PATENT

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re application of: **Mikio ITO et al.**

Serial Number: **Not Yet Assigned**

Filed: **December 19, 2003** **Customer No.: 38834**

For: **RAID APPARATUS AND LOGICAL DEVICE EXPANSION METHOD THEREOF**

CLAIM FOR PRIORITY UNDER 35 U.S.C. 119

Commissioner for Patents
P. O. Box 1450
Alexandria, VA 22313-1450

December 19, 2003

Sir:

The benefit of the filing date of the following prior foreign application is hereby requested for the above-identified application, and the priority provided in 35 U.S.C. 119 is hereby claimed:

Japanese Appln. No. 2002-378284, filed on December 26, 2002

In support of this claim, the requisite certified copy of said original foreign application is filed herewith.

It is requested that the file of this application be marked to indicate that the applicants have complied with the requirements of 35 U.S.C. 119 and that the Patent and Trademark Office kindly acknowledge receipt of said certified copy.

In the event that any fees are due in connection with this paper, please charge our Deposit Account No. 50-2866.

Respectfully submitted,
WESTERMAN, HATTORI, DANIELS & ADRIAN, LLP

William F. Westerman
Reg. No. 29,988

Atty. Docket No.: 032188
Suite 700
1250 Connecticut Avenue, N.W.
Washington, D.C. 20036
Tel: (202) 822-1100
Fax: (202) 822-1111
WFW/yap

日本国特許庁
JAPAN PATENT OFFICE

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office.

出願年月日
Date of Application: ２００２年１２月２６日

出願番号
Application Number: 特願２００２－３７８２８４

[ST. 10/C]: [ＪＰ２００２－３７８２８４]

出願人
Applicant(s): 富士通株式会社
株式会社ＰＦＵ

２００３年１０月２４日

特許庁長官
Commissioner,
Japan Patent Office 今井康夫

出証番号　出証特２００３－３０８８２８１

【書類名】　特許願

【整理番号】　0253075

【提出日】　平成14年12月26日

【あて先】　特許庁長官　殿

【国際特許分類】　G06F　3/06

【発明の名称】　ＲＡＩＤ装置及びその論理デバイス拡張方法

【請求項の数】　10

【発明者】

【住所又は居所】　神奈川県川崎市中原区上小田中４丁目１番１号　富士通株式会社内

【氏名】　伊藤　実希夫

【発明者】

【住所又は居所】　神奈川県川崎市中原区上小田中４丁目１番１号　富士通株式会社内

【氏名】　大黒谷　秀治郎

【発明者】

【住所又は居所】　神奈川県川崎市中原区上小田中４丁目１番１号　富士通株式会社内

【氏名】　池内　和彦

【発明者】

【住所又は居所】　神奈川県川崎市中原区上小田中４丁目１番１号　富士通株式会社内

【氏名】　吉屋　行裕

【発明者】

【住所又は居所】　神奈川県川崎市中原区上小田中４丁目１番１号　富士通株式会社内

【氏名】　大御堂　邦子

【発明者】

【住所又は居所】　石川県河北郡宇ノ気町字宇野気ヌ９８番地の２　株式会社ピーエフユー内

【氏名】　平野　忠司

【特許出願人】

【識別番号】　000005223

【氏名又は名称】　富士通株式会社

【特許出願人】

【識別番号】　000136136

【氏名又は名称】　株式会社ピーエフユー

【代理人】

【識別番号】　100094514

【弁理士】

【氏名又は名称】　林　恒徳

【選任した代理人】

【識別番号】　100094525

【弁理士】

【氏名又は名称】　土井　健二

【手数料の表示】

【予納台帳番号】　030708

【納付金額】　21,000円

【提出物件の目録】

【物件名】　明細書　1

【物件名】　図面　1

【物件名】　要約書　1

【包括委任状番号】　9704944

【包括委任状番号】　0215696

【プルーフの要否】　要

【書類名】明細書

【発明の名称】ＲＡＩＤ装置及びその論理デバイス拡張方法

【特許請求の範囲】

【請求項１】ＲＡＩＤ構成定義に従い、データを分解し、複数の物理デイスク装置に、並列にリード／ライトするＲＡＩＤ装置において、

上位装置からのＩ／Ｏ要求に応じて、前記ＲＡＩＤ構成定義によるＲＬＵマッピングに従い、前記複数の物理デイスク装置をアクセスする制御部と、

少なくとも旧ＲＡＩＤレベルと旧論理デバイス数を定義した旧ＲＡＩＤ構成定義情報と、少なくとも新ＲＡＩＤレベルと新論理デバイス数を定義した新ＲＡＩＤ構成定義情報とを格納するテーブルと、

旧ＲＡＩＤ構成から新ＲＡＩＤ構成に変更するため、データを一時格納するためのキャッシュメモリとを有し、

前記制御部は、前記テーブルの前記旧ＲＡＩＤ構成定義によるＲＬＵマッピングに従い、前記複数の物理デイスク装置からデータを前記キャッシュメモリに読出し、前記テーブルの前記新ＲＡＩＤ構成定義によるＲＬＵマッピングに従い、前記キャッシュメモリに読み出したデータを、前記複数の物理デイスク装置にライトする

ことを特徴とするＲＡＩＤ装置。

【請求項２】前記制御部は、前記旧ＲＡＩＤ構成定義のＲＡＩＤレベルによるＲＬＵマッピングに従い、前記複数の物理デイスク装置からデータを前記キャッシュメモリに読出し、前記新ＲＡＩＤ構成定義のＲＡＩＤレベルによるＲＬＵマッピングに従い、前記キャッシュメモリに読み出したデータを、前記複数の物理デイスク装置にライトするＲＡＩＤレベル変換処理を行う

ことを特徴とする請求項１のＲＡＩＤ装置。

【請求項３】前記制御部は、前記旧ＲＡＩＤ構成定義の前記論理デバイス数によるＲＬＵマッピングに従い、前記複数の物理デイスク装置からデータを前記キャッシュメモリに読出し、前記新ＲＡＩＤ構成定義の前記論理デバイス数によるＲＬＵマッピングに従い、前記キャッシュメモリに読み出したデータを、前記複数の物理デイスク装置にライトする容量増加処理を行う

ことを特徴とする請求項１のＲＡＩＤ装置。

【請求項４】前記制御部は、前記旧ＲＡＩＤ構成から前記新ＲＡＩＤ構成への変換をシーケンシャルに実行し、且つその進捗状況を管理するとともに、前記変換中に、前記上位装置からのＩ／Ｏ要求に対し、変換済み領域かを判定し、変換済み領域に対しては、前記新ＲＡＩＤ構成定義で、前記Ｉ／Ｏ要求を実行し、変換済みでない領域に対しては、前記旧ＲＡＩＤ構成定義で、前記Ｉ／Ｏ要求を実行する

ことを特徴とする請求項１のＲＡＩＤ装置。

【請求項５】前記制御部は、新ＲＡＩＤ構成定義によるＲＬＢＡを上位のＬＢＡに変換した後、前記上位のＬＢＡから前記旧ＲＡＩＤ構成定義によるＲＬＵマッピングに従い、前記複数の物理デイスク装置からデータを前記キャッシュメモリに読出し、前記ＲＬＢＡから前記新ＲＡＩＤ構成定義によるＲＬＵマッピングに従い、前記キャッシュメモリに読み出したデータを、前記複数の物理デイスク装置にライトする

ことを特徴とする請求項１のＲＡＩＤ装置。

【請求項６】ＲＡＩＤ構成定義に従い、データを分解し、複数の物理デイスク装置に、並列にリード／ライトするＲＡＩＤ装置の論理デバイス拡張方法において、

少なくとも旧ＲＡＩＤレベルと旧論理デバイス数を定義した旧ＲＡＩＤ構成定義情報によるＲＬＵマッピングに従い、前記複数の物理デイスク装置からデータを前記キャッシュメモリに読出すステップと、

少なくとも新ＲＡＩＤレベルと新論理デバイス数を定義した新ＲＡＩＤ構成定義情報によるＲＬＵマッピングに従い、前記キャッシュメモリに読み出したデータを、前記複数の物理デイスク装置にライトするステップとからなる

ことを特徴とする論理デバイス拡張方法。

【請求項７】前記読出しステップは、前記旧ＲＡＩＤ構成定義のＲＡＩＤレベルによるＲＬＵマッピングに従い、前記複数の物理デイスク装置からデータを前記キャッシュメモリに読出すステップからなり、

前記ライトステップは、前記新ＲＡＩＤ構成定義のＲＡＩＤレベルによるＲＬ

Ｕマッピングに従い、前記キャッシュメモリに読み出したデータを、前記複数の物理デイスク装置にライトするステップからなる

ことを特徴とする請求項６の論理デバイス拡張方法。

【請求項８】前記読出しステップは、前記旧ＲＡＩＤ構成定義の前記論理デバイス数によるＲＬＵマッピングに従い、前記複数の物理デイスク装置からデータを前記キャッシュメモリに読出すステップからなり、

前記ライトステップは、前記新ＲＡＩＤ構成定義の前記論理デバイス数によるＲＬＵマッピングに従い、前記キャッシュメモリに読み出したデータを、前記複数の物理デイスク装置にライトする容量増加ステップからなる

ことを特徴とする請求項６の論理デバイス拡張方法。

【請求項９】前記旧ＲＡＩＤ構成から前記新ＲＡＩＤ構成への変換をシーケンシャルに実行するその進捗状況を管理するステップと、

前記変換中に、前記上位装置からのＩ／Ｏ要求に対し、変換済み領域かを判定するステップと、

変換済み領域に対しては、前記新ＲＡＩＤ構成定義で、前記Ｉ／Ｏ要求を実行するステップと、

変換済みでない領域に対しては、前記旧ＲＡＩＤ構成定義で、前記Ｉ／Ｏ要求を実行するステップとを更に有する

ことを特徴とする請求項６の論理デバイス拡張方法。

【請求項１０】前記読出しステップは、新ＲＡＩＤ構成定義によるＲＬＢＡを上位のＬＢＡに変換した後、前記上位のＬＢＡから前記旧ＲＡＩＤ構成定義によるＲＬＵマッピングに従い、前記複数の物理デイスク装置からデータを前記キャッシュメモリに読出すステップからなり、

前記ライトステップは、前記ＲＬＢＡから前記新ＲＡＩＤ構成定義によるＲＬＵマッピングに従い、前記キャッシュメモリに読み出したデータを、前記複数の物理デイスク装置にライトするステップからなる

ことを特徴とする請求項６の論理デバイス拡張方法。

【発明の詳細な説明】

【０００１】

【発明の属する技術分野】

本発明は、磁気デイスク等の物理デイスクを使用してデータを冗長管理するＲＡＩＤ装置及びその論理デバイス拡張方法に関し、特に、ＲＡＩＤグループの容量を増加し、又はＲＡＩＤグループに冗長性を付加するＲＡＩＤ装置及びその論理デバイス拡張方法に関する。

【０００２】

【従来の技術】

磁気デイスク、光磁気デイスク、光デイスク等の記憶媒体を利用したストレージ装置では、データ処理装置の要求で、記憶媒体を実アクセスする。データ処理装置が、大容量のデータを使用する場合には、複数のストレージ装置と制御装置とを備えたストレージシステムを利用する。

【０００３】

このようなストレージシステムでは、保存データの信頼性や、装置の信頼性を向上するため、冗長構成を採用している。冗長構成のため、通常、デイスク装置は、ＲＡＩＤ（Redundant Array of Inexpensive（or Independent）Disks）というデイスクの多重化構成を採用する。ＲＡＩＤ機能には、ＲＡＩＤ０，ＲＡＩＤ１，ＲＡＩＤ０＋１，ＲＡＩＤ２，ＲＡＩＤ３，ＲＡＩＤ４，ＲＡＩＤ５が知られている。

【０００４】

このようなＲＡＩＤ構成は、ＲＡＩＤレベルは、システムで固定であった。しかし、冗長度を大きくすると、信頼性は向上するが、性能は低下し、又冗長度を小さくすると、信頼性は低下するが、性能は向上する、という相反した特性がある。ユーザーのシステム構成により、冗長度を設定するが、システム導入後、ユーザーが、冗長度を変更したいとの要求がある。冗長度の変更は、システムを停止すれば、容易に冗長度を変更できる。

【０００５】

しかし、オンラインシステムを構築した場合に、システムを停止することなく、活性状態で、冗長度を変更することが望ましい。従来、パリテイブロックの削減、増加により、冗長度を活性状態で変更する方法が提案されている（例えば、

特許文献１）。

【０００６】

この提案では、ＲＡＩＤ５の構成において、物理デイスク群からデータをキャッシュメモリに読み出し、２パリテイ又は１パリテイから１パリテイ又は０パリテイへの冗長度削減、又は０パリテイ又は１パリテイから１パリテイ又は２パリテイへの冗長度増加を行い、キャッシュメモリから物理デイスク群へ書込むものである。

【０００７】

この冗長度変換処理中、ホストからＩ／Ｏ要求を受けると、冗長度変換処理を中断し、Ｉ／Ｏ要求が、冗長度変更済みの領域に対するものか、冗長度変更前の領域に対するものかを判定し、Ｉ／Ｏ要求を実行するものである。この冗長度削減では、実デイスクの削減を行い、冗長度の増加では、実デイスクの増加を行う必要があった。

【０００８】

【特許文献１】

特開平７－３０６７５８号公報（図２、図７、図８）

【０００９】

【発明が解決しようとする課題】

しかしながら、従来技術は、活性状態での冗長度変更は可能であるものの、パリテイブロックの数を単純に、削減、増加する技術であるため、種々のＲＡＩＤレベルの変更に対応ができないという問題があり、冗長度の変更範囲が限られていた。

【００１０】

更に、活性状態で、ＲＡＩＤレベルを変更せずに、ＲＡＩＤグループの容量を増加するという要求を、実現できないという問題があった。

【００１１】

従って、本発明の目的は、活性状態でのＲＡＩＤレベルの変更範囲を拡大するためのＲＡＩＤ装置及び論理デバイス拡張方法を提供するにある。

【００１２】

又、本発明の他の目的は、物理デイスク数の変更なしに、活性状態でＲＡＩＤレベルを変更するためのＲＡＩＤ装置及び論理デバイス拡張方法を提供するにある。

【００１３】

更に、本発明の他の目的は、活性状態で、ＲＡＩＤレベルを変更せずに、ＲＡＩＤグループの容量を増大するためのＲＡＩＤ装置及び論理デバイス拡張方法を提供するにある。

【００１４】

【課題を解決するための手段】

この目的の達成のため、本発明は、ＲＡＩＤ構成定義に従い、データを分解し、複数の物理デイスク装置に、並列にリード／ライトするＲＡＩＤ装置であり、上位装置からのＩ／Ｏ要求に応じて、前記ＲＡＩＤ構成定義によるＲＬＵマッピングに従い、前記複数の物理デイスク装置をアクセスする制御部と、少なくとも旧ＲＡＩＤレベルと旧論理デバイス数を定義した旧ＲＡＩＤ構成定義情報と、少なくとも新ＲＡＩＤレベルと新論理デバイス数を定義した新ＲＡＩＤ構成定義情報とを格納するテーブルと、旧ＲＡＩＤ構成から新ＲＡＩＤ構成に変更するため、データを一時格納するためのキャッシュメモリとを有し、前記制御部は、前記テーブルの前記旧ＲＡＩＤ構成定義によるＲＬＵマッピングに従い、前記複数の物理デイスク装置からデータを前記キャッシュメモリに読出し、前記テーブルの前記新ＲＡＩＤ構成定義によるＲＬＵマッピングに従い、前記キャッシュメモリに読み出したデータを、前記複数の物理デイスク装置にライトする。

【００１５】

又、本発明の論理デバイス拡張方法は、ＲＡＩＤ構成定義に従い、データを分解し、複数の物理デイスク装置に、並列にリード／ライトするＲＡＩＤ装置の論理デバイス拡張方法において、少なくとも旧ＲＡＩＤレベルと旧論理デバイス数を定義した旧ＲＡＩＤ構成定義情報によるＲＬＵマッピングに従い、前記複数の物理デイスク装置からデータを前記キャッシュメモリに読出すステップと、少なくとも新ＲＡＩＤレベルと新論理デバイス数を定義した新ＲＡＩＤ構成定義情報によるＲＬＵマッピングに従い、前記キャッシュメモリに読み出したデータを、

前記複数の物理デイスク装置にライトするステップとを有する。

【００１６】

本発明では、少なくとも、ＲＡＩＤレベルと論理デバイス数を定義した新旧のＲＡＩＤ構成定義情報を使用し、それぞれによりＲＬＵマッピングして、ＲＡＩＤ構成を変更するため、多様なＲＡＩＤレベルの変換、容量増加を実現できる。

【００１７】

又、本発明では、好ましくは、前記制御部は、前記旧ＲＡＩＤ構成定義のＲＡＩＤレベルによるＲＬＵマッピングに従い、前記複数の物理デイスク装置からデータを前記キャッシュメモリに読出し、前記新ＲＡＩＤ構成定義のＲＡＩＤレベルによるＲＬＵマッピングに従い、前記キャッシュメモリに読み出したデータを、前記複数の物理デイスク装置にライトするＲＡＩＤレベル変換処理を行う。

【００１８】

又、本発明では、好ましくは、前記制御部は、前記旧ＲＡＩＤ構成定義の前記論理デバイス数によるＲＬＵマッピングに従い、前記複数の物理デイスク装置からデータを前記キャッシュメモリに読出し、前記新ＲＡＩＤ構成定義の前記論理デバイス数によるＲＬＵマッピングに従い、前記キャッシュメモリに読み出したデータを、前記複数の物理デイスク装置にライトする容量増加処理を行う。

【００１９】

又、本発明では、好ましくは、前記制御部は、前記旧ＲＡＩＤ構成から前記新ＲＡＩＤ構成への変換をシーケンシャルに実行し、且つその進捗状況を管理するとともに、前記変換中に、前記上位装置からのＩ／Ｏ要求に対し、変換済み領域かを判定し、変換済み領域に対しては、前記新ＲＡＩＤ構成定義で、前記Ｉ／Ｏ要求を実行し、変換済みでない領域に対しては、前記旧ＲＡＩＤ構成定義で、前記Ｉ／Ｏ要求を実行する。

【００２０】

又、本発明では、好ましくは、前記制御部は、新ＲＡＩＤ構成定義によるＲＬＢＡを上位のＬＢＡに変換した後、前記上位のＬＢＡから前記旧ＲＡＩＤ構成定義によるＲＬＵマッピングに従い、前記複数の物理デイスク装置からデータを前記キャッシュメモリに読出し、前記ＲＬＢＡから前記新ＲＡＩＤ構成定義による

ＲＬＵマッピングに従い、前記キャッシュメモリに読み出したデータを、前記複数の物理デイスク装置にライトする。

【００２１】

又、本発明では、好ましくは、前記制御部は、前記旧ＲＡＩＤ構成から前記新ＲＡＩＤ構成への変換後、前記旧ＲＡＩＤ構成定義を前記テーブルから削除する。

【００２２】

又、本発明では、好ましくは、前記制御部は、指示された新ＲＡＩＤ構成定義のパラメータと前記旧ＲＡＩＤ構成定義とに従い、前記新ＲＡＩＤ構成定義を、前記テーブルに作成する。

【００２３】

又、本発明では、好ましくは、前記制御部は、変換領域に対応する前記キャッシュメモリの領域を獲得した後、前記旧ＲＡＩＤ構成から前記新ＲＡＩＤ構成への変換をシーケンシャルに実行する。

【００２４】

又、本発明では、好ましくは、前記制御部は、前記変換領域に対応する前記キャッシュメモリの領域が獲得できない時は、前記変換処理を複数回に分けて実行する。

【００２５】

又、本発明では、好ましくは、前記制御部は、前記ＲＡＩＤ構成のストライプに対応したストライプデプス、ストライプサイズに応じて、前記ＲＬＵマッピングする。

【００２６】

【発明の実施の形態】

以下、本発明の実施の形態を、ストレージシステム、ＲＡＩＤ構成、ＬＤＥ，ＬＤＥ全体処理、ＬＤＥ詳細処理、容量増加処理／ＲＡＩＤレベル変換処理、他の実施の形態の順で説明する。

【００２７】

［ストレージシステム］

図１は、本発明の一実施の形態のストレージシステムの構成図であり、磁気デイスクを使用したＲＡＩＤ（Redundant Arrays of Inexpensive Disk）システムを示す。

【００２８】

図１に示すように、ストレージシステムは、一対の磁気デイスクコントローラ（以下、コントローラという）１、２と、この一対のコントローラ１、２にラインｌ１，ｌ２で接続された多数の磁気デイスク装置５０－１～５０－ｍ、５２－１～５２－ｎとからなる。

【００２９】

コントローラ１、２は、直接又はネットワーク機器を介し、ホストやサーバーに接続され、ホストやサーバーの大量のデータを、ＲＡＩＤデイスクドライブ（磁気デイスク装置）へ高速かつ、ランダムに読み書きが出来るシステムである。一対のコントローラ１、２は、同一の構成を有し、ＣＡ（Channel Adapter）１１、１２、２１、２２と、ＣＭ（Centralized Module）１０、１５～１９、２０、２５～２９と、ＤＡ（Device Adapter）１３、１４、２３、２４のファンクションモジュールによって構成されている。

【００３０】

ＣＡ（Channel Adapter）１１、１２、２１、２２は、ホストを結ぶホスト・インタフェースの制御をつかさどる回路であり、例えば、ファイバーチャネル回路（ＦＣ）とＤＭＡ（Direct Memory Access）回路等で構成される。ＤＡ（Device Adapter）１３、１４、２３、２４は、デイスクデバイス５０－１～５０－ｍ、５２－１～５２－ｍを制御するため、デイスクデバイスとコマンド、データのやり取りを行う回路であり、例えば、ファイバーチャネル回路（ＦＣ）とＤＭＡ回路等で構成される。

【００３１】

ＣＭ（Centralized Module）は、ＣＰＵ１０，２０と、ブリッジ回路１７、２７と、メモリー（ＲＡＭ）１５、２５と、フラッシュメモリ１９，２９と、ＩＯブリッジ回路１８，２８とを有する。メモリー１５，２５は、バッテリーでバックアップされ、その一部が、キャッシュメモリ１６，２６として使用される。

【００３２】

ＣＰＵ１０，２０は、ブリッジ回路１７を介し、メモリー１５，２５、フラッシュメモリ１９，２９、ＩＯブリッジ回路１８，２８に接続される。このメモリー１５，２５は、ＣＰＵ１０，２０のワーク領域に使用され、フラッシュメモリ１９，２９は、ＣＰＵ１０，２０が実行するプログラムを格納する。

【００３３】

このプログラムとして、ＯＳ，ＢＩＯＳ（Basic Input/Output System），ファイルアクセスプログラム（リード／ライトプログラム）、ＲＡＩＤ管理プログラム等の制御プログラム（モジュール）を格納する。ＣＰＵ１０，２０は、このプログラムを実行し、後述するように、リード／ライト処理、ＲＡＩＤ管理処理等を実行する。

【００３４】

ＰＣＩ（Personal Computer Interface）バス３１は、ＣＡ１１，１２，２１，２２と、ＤＡ１３，１４，２３，２４とを接続するとともに、ＩＯブリッジ回路１８を介し、ＣＰＵ１０，２０、メモリー１５，２５を接続する。更に、ＰＣＩバス３１には、ＰＣＩ－ノードリンクブリッジ回路３０，４０が接続される。コントローラ１のＰＣＩ－ノードリンクブリッジ回路３０は、コントローラ２のＰＣＩ－ノードリンクブリッジ回路４０と接続され、コントローラ１，２間のコマンド、データの交信を行う。

【００３５】

図１では、例えば、コントローラ１は、デイスク装置５０－１～５０－ｍを担当し、コントローラ２は、デイスク装置５２－１～５２－ｎを担当する。又、デイスク装置５０－１～５０－ｍと、５２－１～５２－ｎとが、ＲＡＩＤ５の構成を有する。

【００３６】

キャッシュメモリ１６，２６は、各々、担当するデイスク装置のデータの一部を格納し、ホストからのライトデータを格納する。ＣＰＵ１０，２０は、ＣＡ１１，１２，２１，２２を介しホストからのリード要求を受けて、キャッシュメモリ１６，２６を参照し、物理デイスクへのアクセスが必要かを判定し、必要であ

れば、デイスクアクセス要求をＤＡ１３，１４，２３，２４に要求する。又、ＣＰＵ１０，２０は、ホストからのライト要求を受けて、ライトデータをキャッシュメモリ１６，２６に書込み、且つ内部でスケジュールされるライトバック等をＤＡ１３，１４，２３，２４に要求する。

【００３７】

［ＲＡＩＤ構成］

図２は、ＣＰＵ１０が実行するタスクと制御モジュールの構成図である。図２に示すように、ＣＭ－ＣＡドライバ３２は、ＣＡ１１，１２をドライブするドライバである。Ｂａｓｉｃ　Ｔａｓｋ３３は、ＣＰＵ１０が実行する基本タスクであり、資源管理を行うリソーススレッド（リソース制御モジュール）３５、コピー処理を行うコピースレッド（コピー制御モジュール）３６、キャッシュメモリ１６の制御を行うキャッシュスレッド（キャッシュメモリ制御モジュール）３７、ＲＡＩＤ構成制御を行うＲＡＩＤスレッド（ＲＡＩＤ制御モジュール）３９、メインテナンスエージェント（メインテナンス制御モジュール）３４、ＯＶＳＭスレッド（ＯＶＳＭ制御モジュール）３８とを有する。

【００３８】

ＯＶＳＭスレッド３８は、後述するように、クイックフォーマット（ＱＦ）や論理デバイス拡張（ＬＤＥ）のスケジューリング、ＲＡＩＤスレッド３９への論理フォーマット（ＬＦ）依頼、ＬＤＥ進捗管理を行う。メインテナンスエージェント３４は、ＯＶＳＭスレッド３８への各種の通知を行う。

【００３９】

ＣＭ－ＤＡドライバ４２は、ＣＭ１３，１４をドライブするドライバである。ＣＭ－ＣＭドライバ４３は、前述のＰＣＩ－ノードブリッジ回路３０をドライブして、ＣＭ間の通信を行う。

【００４０】

保守ＰＣ（パーソナルコンピュータ）３は、図１では、図示していないが、ブリッジ回路１７のＬＡＮポートに接続され、ストレージシステムの保守を行うものであり、本発明では、ＬＤＥの要求、ＬＤＥ進捗報告の表示等を行う。ＨＴＴＰドライバ４４は、保守ＰＣ３とのHTTP（Hyper　Text　Transfer　Protocol）

のドライバである。

【００４１】

ＣＧＩ（Computer Graphic Interface）タスク４５は、保守ＰＣ３からのＬＤＥ要求の受け付けと、保守ＰＣ３へのＬＤＥ進捗報告を行う。メインテナンスドライバ４６は、メインテナンスのためのメインテナンスＰＣスレッド４７とシステム制御スレッド４８のドライバである。

【００４２】

メインテナンスＰＣスレッド（制御モジュール）４７は、保守ＰＣ３からの要求に応じ、システム制御スレッド４８を起動する。システム制御スレッド（制御モジュール）４８は、ＱＦ／ＬＤＥ開始時のＲＬＵ(Raid Logical Unit)テーブル（図３、図１４で後述する）の変更、ＯＶＳＭスレッド３８へのＱＦ／ＬＤＥ開始指示、ＱＦ／ＬＤＥ終了時のＲＬＵテーブル（図３、図１４にて後述する）の変更を行う。

【００４３】

本発明では、後述するように、保守ＰＣ３のＬＤＥ要求により、ＯＶＳＭスレッド３８が、キャッシュスレッド３７、ＲＡＩＤスレッド３９を制御し、ＬＤＥ処理を実行する。

【００４４】

図３は、ＲＡＩＤ論理空間の説明図、図４は、ＲＡＩＤ５の説明図、図５は、ＲＡＩＤ５のアドレス計算フロー図、図６は、ＲＡＩＤ０＋１の説明図、図７は、ＲＡＩＤ０＋１のアドレス計算フロー図である。

【００４５】

図３に示すように、ホストから見たＲＡＩＤ論理空間は、ホストの論理空間であるＯＬＵ（ホスト論理ユニット）と、ＲＡＩＤグループの論理空間であるＲＬＵ（ＲＡＩＤ論理ユニット）と、ＲＡＩＤグループを構成するデバイスの論理空間であるＤＬＵ（デバイス論理ユニット）と、物理デイスクの論理空間であるＰＬＵ（物理論理ユニット）との層構造で示される。

【００４６】

ＲＡＩＤ構成では、ＲＡＩＤ論理空間は、ＯＬＵテーブル７０の開始ＲＬＢＡ

（ＲＡＩＤ論理ブロックアドレス）で、ＯＬＵと関連づけられ、ＲＡＩＤ空間は、ＲＬＵテーブル７２で定義される。ＲＬＵテーブル７２は、ＲＡＩＤレベル、構成デイスク数、ＲＡＩＤストライプデプス、ＲＡＩＤストライプサイズ、対応ＤＬＵ番号を格納する。

【００４７】

ＤＬＵ空間は、ＤＬＵテーブル７４で定義される。ＤＬＵテーブル７４は、構成デイスク数、ＲＡＩＤストライプデプス、ＲＡＩＤストライプサイズ、対応ＰＬＵ番号を格納する。ＤＬＵ空間及びＤＬＵテーブル７４は、ミラーリングで使用される。ＰＬＵ空間は、ＰＬＵテーブル７６で定義される。ＰＬＵテーブル７６は、開始ＰＬＢＡ（物理論理ブロックアドレス）を格納する。

【００４８】

具体的に説明する。図４に示すように、ＲＡＩＤ５（３＋１）の場合には、ＲＬＵ＝ＤＬＵであり、ＲＬＵテーブル７２は、ＲＡＩＤレベル＝ＲＡＩＤ５，構成デイスク数＝４、対応ＤＬＵ番号＝ＰＬＵ番号（０～３）となる。又、ＲＡＩＤ空間は、構成デイスクで、ストライピングされており、ＲＡＩＤ空間を、構成デイスク番号とストライプ番号とでマッピングする。この升目をストリップ（Ｓｔｒｉｐ）といい、Strip番号が付与される。このStripのサイズは、Strip Depth（又はStripe Depth）で定義され、１ストライプのサイズは、Stripe Sizeで定義される。

【００４９】

従って、図５で説明するように、Ｒ（ＲＡＩＤグループ）ＬＢＡは、構成デイスク数、Strip Depth、Stripe Sizeで、ＰＬＵＬＢＡ，メンバーデイスクの順番に変換できる。

【００５０】

（Ｓ１０）ホストＬＢＡ（論理ブロックアドレス）を、ホストＬＢＡにＯＬＵテーブル７０の開始ＲＬＢＡを加算して、ＲＬＢＡを求める。

【００５１】

（Ｓ１２）ストリップ内のブロックカウントを、ＲＬＵＬＢＡ／Ｓｔｒｉｐ Ｓｉｚｅ（Stripe Depth）の余りで計算する。

【００５２】

（Ｓ１４）Strip番号を、ＲＬＵＬＢＡ／Ｓｔｒｉｐ　Ｄｅｐｔｈから計算する。

【００５３】

（Ｓ１６）Ｓｔｒｉｐｅ番号を、ＲＬＵＬＢＡ／Ｓｔｒｉｐｅ　Ｓｉｚｅから計算する。

【００５４】

（Ｓ１８）メンバーデイスクの順番を、Ｓｔｒｉｐ番号／メンバーデイスクの数の余りから計算する。

【００５５】

（Ｓ２０）物理デイスク（ＰＬＵ）のＬＢＡを、（Ｓｔｒｉｐｅ番号×Strip Size）＋ストリップ内のブロックカウントから計算する。

【００５６】

これにより、メンバーデイスクの順番（ＰＬＵ番号）と、ＰＬＵＬＢＡとから、ＰＬＵテーブル７６を用いて、実ブロックアドレスを計算する。

【００５７】

同様に、図６に示すように、ＲＡＩＤ０＋１（４＋４）の場合には、ＲＬＵ＝ＤＬＵであり、ＲＬＵテーブル７２は、ＲＡＩＤレベル＝ＲＡＩＤ０＋１，構成デイスク数＝４（ＤＬＵ）、対応ＤＬＵ番号＝０～３となる。又、ＲＡＩＤ空間は、ＤＬＵ構成デイスクで、ストライピングされており、ＲＡＩＤ空間を、ＤＬＵ構成デイスク番号とストライプ番号とでマッピングする。この升目をストリップ（Ｓｔｒｉｐ）といい、Strip番号が付与される。このStripのサイズは、Strip Depth（又はStripe Depth）で定義され、１ストライプのサイズは、Stripe Sizeで定義される。

【００５８】

従って、図７で説明するように、Ｒ（ＲＡＩＤグループ）ＬＢＡは、構成デイスク数、Strip Depth、Stripe Sizeで、ＰＬＵＬＢＡ，メンバーデイスクの順番に変換できる。

【００５９】

（Ｓ２２）ホストＬＢＡ（論理ブロックアドレス）を、ホストＬＢＡにＯＬＵテーブル７０の開始ＲＬＢＡを加算して、ＲＬＢＡを求める。

【００６０】

（Ｓ２４）ストリップ内のブロックカウントを、ＲＬＵＬＢＡ／Ｓｔｒｉｐｅ Ｄｅｐｔｈ）の余りで計算する。

【００６１】

（Ｓ２６）Ｓｔｒｉｐｅ番号を、ＲＬＵＬＢＡ／Ｓｔｒｉｐｅ　Ｓｉｚｅから計算する。

【００６２】

（Ｓ２８）メンバーデイスクの順番を、（ＲＬＵＬＢＡ／Ｓｔｒｉｐｅ　Ｓｉｚｅの余り）／Ｓｔｒｉｐ　Ｓｉｚｅから計算する。

【００６３】

（Ｓ３０）物理デイスク（ＰＬＵ）のＬＢＡ（＝ＤＬＵのデイスク）を、（Ｓｔｒｉｐｅ番号×Strip Size）＋ストリップ内のブロックカウントから計算する。

【００６４】

これにより、メンバーデイスクの順番（ＤＬＵ番号）と、ＰＬＵＬＢＡとから、ＰＬＵテーブル７６を用いて、実ブロックアドレスを計算する。

【００６５】

［ＬＤＥ］

次に、ＬＤＥを説明する。Logical Device Expansion(LDE)は、（１）デイスク装置を増設したり、ＲＡＩＤレベルを変換することで、ＲＡＩＤグループの容量を増加させる、（２）ＲＡＩＤレベルを変換することで、ＲＡＩＤグループに冗長性を付加する，機能である。ＲＡＩＤ容量を拡張する方法として、ＲＡＩＤグループへの新規デイスク装置増設と、ＲＡＩＤレベル変換とがある。

【００６６】

図８及び図９は、ＲＡＩＤグループへの新規デイスク装置増設（容量増加）の説明図である。容量増加は、ＲＡＩＤグループにユーザデータを保持したまま、新規デイスク装置(New Disk)を追加して、ＲＡＩＤグループの容量を増やす機能

である。

【００６７】

例えば、図８に示すように、ＲＡＩＤ５（４＋１）に、新規デイスクを1台追加して、ＲＡＩＤ５（５＋１）にする。追加デイスクの容量が、３６ＧＢなら、ＲＡＩＤグループ１４４ＧＢを，１８０ＧＢに容量を増やす。

【００６８】

また、図９に示すように、全ユーザデータを、容量の大きな新規デイスク装置で構成されたＲＡＩＤグループに遷移することでも，ＲＡＩＤ容量の拡張を実現できる。例えば、１８ＧＢのデイスク装置で構成されたＲＡＩＤ５（４＋１）容量７２ＧＢのユーザデータを、新規デイスク装置（３６ＧＢ）５台を用いて構成されたＲＡＩＤ５（４＋１）容量１４４ＧＢへ遷移し、ＲＡＩＤ容量を増やす。

【００６９】

次に、ＲＡＩＤレベル変換を、図１０乃至図１２は、ＲＡＩＤレベル変換の説明図である。ＲＡＩＤレベル変換は、ユーザデータをＲＡＩＤグループに保持したまま、ＲＡＩＤレベルの変更を行う機能である。

【００７０】

例えば、図１０に示すように、複数台（４台）の新規デイスク装置を追加し、ＲＡＩＤ５（４＋１）からＲＡＩＤ０＋１（４＋４）への変更を行う。

【００７１】

また、ＲＡＩＤグループを構成している既存デイスク装置を用いず、新規デイスク装置だけを用いてのＲＡＩＤレベルの変更も出来る。例えば、図１１に示すように、１８ＧＢデイスク装置４台構成のＲＡＩＤ５（３＋１）を、７３ＧＢデイスク装置２台構成のＲＡＩＤ１へ変更する。

【００７２】

本実施の形態では、可能なＲＡＩＤレベル変換は、図１２に示すように、ＲＡＩＤ０からＲＡＩＤ０＋１，ＲＡＩＤ１、ＲＡＩＤ５、ＲＡＩＤ０＋１とＲＡＩＤ１、ＲＡＩＤ５と相互間、ＲＡＩＤ１とＲＡＩＤ５との相互間である。ＲＡＩＤ０への変換は、冗長性を失うことになるため実施しない。

【００７３】

このＬＤＥ実行中、ＬＤＥを担当するＣＭを切り替えることができる。又、ＬＤＥを、パワーオフ／オンを跨いで継続され、停電/復電を跨いで継続される。更に、ＣＭ活性増設・活性交換は実行できる。

【００７４】

但し、ＲＡＩＤ１へのＲＡＩＤレベル変換は、既存デイスクを使用せずに新規デイスクのみで作成し、その新規デイスクは、既存デイスクよりも容量が大きくないことが必要である。又、ＲＡＩＤレベル変換について、ＲＡＩＤ０への変換と容量が減少する変換は行わない。

【００７５】

［ＬＤＥ全体処理］

次に、図１３及び図１４により、図２の構成のＬＤＥ時のＣＧＩタスク４５からの流れを説明する。先ず、ＣＧＩタスク４５より保守タスク４７へＬＤＥ設定通知を行う。このときのパラメタは、図９で説明するように、ＬＤＥ内容（追加、レベル変換等）、ＲＬＵＮ，ＲＡＩＤレベル、新規構成デイスクである。保守タスク４７は、装置状態がＬＤＥ実行可能かどうかをチェックする。更に、保守タスク４７は、システム制御モジュール４８に対して、設定パラメタがＬＤＥ実行可能な条件（図１２参照）を満たしているかどうかのチェックを要求する。

【００７６】

ＣＧＩタスク４５は、保守タスク４７からＬＤＥ実行可能な応答を受けて、保守タスク４７へＬＤＥ設定を通知する。これに応じて、保守タスク４７は、システム制御４８にＬＤＥ設定通知を行う。システム制御４８は，キャッシュ制御３７に対して、WTTH（Write　Through）で動作するよう通知する。キャッシュ制御３７は、以降、WTTHで動作する。

【００７７】

次に、保守タスク４７は、構成変更を行うためシステム制御４８に対して、Suspendを通知する。システム制御４８は、それを受けて保守エージェント３４へSuspend要求を出す。保守エージェント３４は、ＢａｃｋＥｎｄ（Cacheモジュール３７、ＲＡＩＤモジュール３９）に対してSuspend要求を出す。これにより、対象ＲＬＵを含めてすべてのＩ／Ｏを一時的に抑制する。

【００７８】

更に、システム制御４８は、Suspend処理が完了すると、ＯＶＳＭ３８に対して構成変更を要求する。ＯＶＳＭ３８は、ＬＤＥを開始するための構成を作成する。図１４に示すように、現在のＲＡＩＤ構成定義、即ち、ＲＬＵ７２，ＤＬＵ７４，ＰＬＵ７６を、旧構成のテンポラリＲＡＩＤ定義８０にコピーし、前述のＣＧＩ４５から通知されたＬＤＥ内容（追加、レベル変換等）、ＲＬＵＮ，ＲＡＩＤレベル、新規構成デイスクから新構成のＲＡＩＤ定義８２、即ち、ＲＬＵ７２，ＤＬＵ７４，ＰＬＵ７６を作成する。ＯＶＳＭ３８は、変更内容を各モジュールに対して配信する。

【００７９】

この処理を終了後、システム制御４８は、保守エージェント３４に対して、ＬＤＥ実行要求を出す。それを受けて保守エージェント３４は，ＯＶＳＭ３８に対して、ＬＤＥ実行要求を出す。ＯＶＳＭ３８は、ＬＤＥ初期処理を実行し、保守エージェント３４へ応答を返す。その後、ＯＶＳＭ３８は、図１４のＬＤＥ進捗状況を作成し、ＬＤＥ処理を実行する。保守エージェント３４は、応答をシステム制御４８へ返す。

【００８０】

この後、システム制御４８は、保守エージェント３４に対して、Resumeを通知する。保守エージェント３４は、ＢａｃｋＥｎｄへResume要求を出す。これによって、Ｉ／Ｏを再開する。システム制御４８は、キャッシュ制御３７に対して、WB(Write　Back)modeへ戻るように通知する。システム制御３７は、ＣＧＩタスク４５に対して、ＬＤＥが開始されたことを通知する。

【００８１】

次に、ＣＧＩタスク４５は、保守タスク４７経由で、ＯＶＳＭ３８に対して、ＬＤＥ進捗情報獲得要求を出し、進捗情報を獲得する。ＯＶＳＭ３８は、ＬＤＥ処理が完了すると、ＣＶMmodeのＬＤＥFlagをOFFにする等の構成変更を行い、旧構成情報を削除する。ＯＶＳＭ３８は、ＢａｃｋＥｎｄに対して、構成を配信する。ＯＶＳＭ３８は、ＬＤＥ処理の後処理を行う。

【００８２】

［ＬＤＥ詳細処理］

次に、ＬＤＥ処理を、図１５乃至図２３で説明する。図１５は、ＬＤＥ制御モジュールの構成図、図１６は、図１５のＬＤＥ処理の説明図、図１７は、図１５のＬＤＥ起動時の処理フロー図、図１８は、図１５のキャッシュメモリ獲得処理フロー図、図１９は、図１５のシーケンシャルＬＥＤ処理の説明図、図２０は、図１５の進捗状況管理によるＩ／Ｏ競合時の処理フロー図である。

【００８３】

図１５に示すように、ＯＶＳＭ３８は、Ｅｘｐａｎｓｉｏｎ　Ｂｌｏｃｋを決定し、キャッシュメモリ１６の領域を獲得する処理９０を実行し、ＲＡＩＤスレッド３９に、ＬＤＥのリード／ライトを依頼する。又、ＲＡＩＤスレッド３９のＬＤＥの進捗状況を管理する処理９２を行う。

【００８４】

ＲＡＩＤスレッド３９は、キャッシュスレッド３７と共同して、Expansion実施前のＲＡＩＤグループ構成（図１４の旧構成テーブル８０）で、デイスク装置５０へのリード処理を行い、キャッシュメモリ１６の獲得領域にステージング処理９４を行う。

【００８５】

次に、ＲＡＩＤスレッド３９は、Expansion実施後のＲＡＩＤグループ構成（図１４の新構成テーブル８２）で、キャッシュメモリ１６からデイスク装置５０へのライト（ライトバック）処理９６を行う。

【００８６】

この時、ＬＤＥでは、リード／ライト処理時のアドレスが異なる。この処理を、増設の場合には、新デイスク装置を含めた領域まで実施したところで、Expansionが完了する。このＬＤＥの実施にあたり、図１４で説明したように、Expansion前(旧構成)のＲＡＩＤグループ構成情報は、テンポラリＲＡＩＤグループ定義情報８０として引継ぎ、Expansion処理の終了時に削除する。

【００８７】

図１６は、ＲＡＩＤにおけるExpansion処理（容量拡張処理）の説明図であり、４台のデイスク装置に新規デイスク装置１台を追加する例を示す。新構成のメ

ンバーデイスク装置数（ここでは、５台）分のストライプ（図の枠内）を、旧構成のデイスク装置（４台）からリードする。次に、旧構成分のメンバーデイスク装置数分のストライプを、新構成のデイスク装置５台にライトする。図１６では、ＲＡＩＤ０＋１において、（４＋４）構成から、（５＋５）構成に変更する例を示し、旧構成（４＋４）から５つのストライプをリードし、新構成（５＋５）に４つのストライプをライトする。

【００８８】

次に、図１７により、保守エージェントよりＬＤＥ開始要求を受けてからのＯＶＳＭの制御シーケンスフローを説明する。保守エージェント３４は、ＯＶＳＭ３８に対し、ＬＤＥの起動を指示する。このときのパラメタは、ＬＤＥ内容（追加、レベル変換等），RLUN，RAIDレベル，新規構成デイスクである。

【００８９】

ＯＶＳＭ３８は、受け取ったパラメタから、構成チエックを行う。即ち、ＲＬＵＮ８２とＴ－ＲＬＵＮ８０の比較、ＬＤＥタイプ等である。そして、ＯＶＳＭ３８は、シーケンシャルExpansion用のＡＣＢ（Ａｃｃｅｓｓ　Ｃｏｎｔｒｏｌ　Ｂｌｏｃｋ）を獲得する。尚、ＡＣＢの数だけ平行動作が可能となる。且つExpansionを行うにあたり、デイスク装置上のデータを移動するために、キャッシュメモリ１６を獲得する（図１８にて詳細に説明する）。

【００９０】

更に、ＯＶＳＭ３８は、他系ＣＭ（図１のコントローラ２）へ制御テーブル及びキャッシュ域獲得を要求する。他系ＣＭは，制御テーブル及びキャッシュ域の獲得を行う。これにより、制御テーブル及びキャッシュ域の二重化が可能となる。ＯＶＳＭ３８は、他ＣＭからキャッシュ域獲得の応答を受け取ると、保守エージェント３４へＬＤＥ起動応答を通知する。

【００９１】

そして、ＯＶＳＭ３８は、シーケンシャルなＥｘｐａｎｓｉｏｎ開始をＲＡＩＤスレッド３９に通知する。即ち、ＯＶＳＭ３８は、一回のExpansion処理領域とＩ／Ｏの排他を行う（図２０で詳述する）。そして、ＲＡＩＤスレッド３９に、Expansion実施中領域について、テンポラリRLUN(Expansion前構成)８０でのリ

ード要求を行う。

【００９２】

ＯＶＳＭ３８は、他系ＣＭとＬＤＥステータス情報の二重化を行う。そして、ＯＶＳＭ３８は、ＲＡＩＤスレッド３９に、Expansion実施中領域について、RLUN(Expansion後構成)８２でライト要求を行う。ＯＶＳＭ３８は、Expansion実施中領域の進捗情報を更新して、二重化されている管理テーブルの進捗情報更新処理を行うために、他系ＣＭへ通信を実施する。ＯＶＳＭ３８は、一回のExpansion処理領域とＩ／Ｏとの排他を解除する。

【００９３】

次に、図１５、図１７で説明したキャッシュメモリ獲得処理９０を、図１８及び図１９で説明する。図１５で説明したように、Ｅｘｐａｎｓｉｏｎ処理中は、デイスク上のデータを移動させるので、一時的にキャッシュメモリ１６にデータを置く。そのため、Sequential Expansion実施前にキャッシュメモリ１６を獲得しておく。キャッシュメモリ１６は、一度に獲得できる大きさが決まっているので、必要容量を獲得するためには、ＡＣＢを必要数獲得してから実施する。

【００９４】

図１８に示すように、１回の処理サイズを決定し、その処理サイズが、一度に獲得できるキャッシュメモリの限界サイズかを判定する。獲得可能なサイズなら、キャッシュスレッド３７に１回だけ、獲得要求を発する。一方、獲得可能でないなら、複数回獲得できるかを判定し、複数回獲得できれば、必要回数要求し、複数回獲得できなければ、１回だけ要求し、複数回使いまわす。即ち、Expansion実施領域のＬＤＥを、一度に実施せずに複数回に分けて実施する。または、必要ＡＣＢ等を獲得してから、平行処理を行うことにする。

【００９５】

例えば、図１８のようなExpansion処理において、一回のExpansion領域を複数回のリード／ライトを実施していき、次の領域へ進む。具体的に説明する。

【００９６】

(1)リード処理をExpansion後のストライプ分(３ストライプ)行う。

【００９７】

(2)(1)で読み込んだ領域について、ライト処理を行う。

【００９８】

(3)リード処理をExpansion後のストライプ分(1ストライプ)行う。

【００９９】

(4)(3)で読み込んだ領域について、ライト処理を行う。

【０１００】

図１８では、ＲＡＩＤ０＋１の（４＋４）からＲＡＩＤ０＋１の（５＋５）への構成変更の例を示し、Expansionブロック数＝旧構成MemberDisk数×新構成MemberDisk数＝4×5＝２０ブロック（ストリップ）であり、ロックするストライプ数＝キャッシュメモリサイズ÷(新構成MemberDisk数×60KB)＝１MB÷(5×60KB)＝３ストライプとなる。又、Expansion実行数＝Expansionブロック数÷(ロックするストライプ数×新構成MemberDisk数)＝20÷(3×5)＝２回であり、最後のExpansion実行ブロック数＝Expansionブロック数%・(ロックするストライプ数 × 新構成MemberDisk数)＝20%・(3×5)＝５ブロックとなる。但し、キャッシュ域は、１MB，Stripは、６０ＫＢとする。

【０１０１】

又、ＲＡＩＤ構成の（１４＋１）から（１５＋１）へのExpansionの場合には、Expansionブロック数＝14×15=210ブロック、ロックするストライプ数＝１MB÷(15×60KB)＝１ストライプ、Expansion実行数＝210÷(1×15)＝１４回、最後のExpansion実行ブロック数=210%・(1×15)＝１５ブロックとなる。

【０１０２】

次に、図２０により、ＯＶＳＭ３８のＩ／Ｏ競合処理を説明する。

【０１０３】

（Ｓ４０）ＯＶＳＭ３８は、ＲＬＵを、ＲＬＢＡにより、Expansion実施済みか、実施中か、未実施かを判定する進捗管理機能９２（図１５参照）により、ホストＩ／Ｏ要求のＲＬＢＡが、Expansion実施領域か未実施領域かの判断を行う。実施済み領域の場合、ＲＡＩＤスレッド３９へ新構成情報でＩ／Ｏ要求を行い、ホストＩ／Ｏ要求を実行する。

【０１０４】

（Ｓ４２）未実施の場合には、Expansion実施中領域かの判断を行う。実施中の場合には、Expansion処理が終了するまで、ホストＩ／Ｏ要求を待たせ、終了後に，ＲＡＩＤスレッド３９へＩ／Ｏ要求を行う。一方、実施中領域でない場合には、Expansion未実施領域のため、ＲＬＵ構成情報をテンポラリＲＬＵ構成情報(旧構成)８０に切り替えてから、ＲＡＩＤスレッド３９へＩ／Ｏ要求を行う。

【０１０５】

次に、図１５で説明したリード処理、ライト処理を、図２１乃至図２３で説明する。図２１は、リード処理のリードアドレス生成処理フロー図である。

【０１０６】

（Ｓ５０）先ず、新ＲＡＩＤ構成定義８２で、新ＲＬＢＡ（ＲＬＵＢＬＡ）を生成する。

【０１０７】

（Ｓ５２）図３で説明したＯＬＵテーブル７０を使用して、ＯＬＢＡに逆変換する。

【０１０８】

（Ｓ５４）ＯＬＢＡを旧構成のＯＬＵテーブル７０で、旧ＲＬＢＡに変換する。

【０１０９】

（Ｓ５６）旧ＲＬＢＡを、旧ＲＡＩＤ構成定義のＲＬＵテーブル７２、ＤＬＵテーブル７４を使用して、旧ＤＬＢＡに変換する。

【０１１０】

（Ｓ５８）旧ＲＡＩＤ構成定義のＰＬＵテーブル７６を使用して、旧ＤＬＢＡをＰＬＢＡに変換し、リードブロックアドレスを得る。

【０１１１】

このリードブロックアドレスで、デイスク装置をリードし、データをキャッシュメモリ１６にリード（ステージング）する。

【０１１２】

図２２は、ライト処理のライトアドレス生成処理フロー図である。

【０１１３】

（Ｓ６０）先ず、新ＲＡＩＤ構成定義８２で、新ＲＬＢＡ（ＲＬＵＢＬＡ）を生成する。

【０１１４】

（Ｓ６２）図３で説明した新ＲＡＩＤ構成定義のＲＬＵテーブル７２、ＤＬＵテーブル７４を使用して、新ＤＬＢＡに変換する。

【０１１５】

（Ｓ６４）新ＲＡＩＤ構成定義のＰＬＵテーブル７６を使用して、新ＤＬＢＡをＰＬＢＡに変換し、ライトブロックアドレスを得る。

【０１１６】

このライトブロックアドレスで、キャッシュメモリ１６のデータを、デイスク装置にライトバックする。

【０１１７】

図２３は、この関係を示す説明図であり、新ＲＡＩＤ構成定義での新ＲＬＢＡを上位のホストＬＢＡに直し、このホストＬＢＡから旧ＲＡＩＤ定義８０で、旧ＲＡＩＤ定義でのリードＰＬＢＡに変換し、リードする。又、新ＰＬＢＡを、新ＲＡＩＤ構成定義８２で、新ＲＡＩＤ定義でのライトＰＬＢＡに変換し、リードしたデータをライトする。

【０１１８】

このようにすると、図４乃至図７で説明した既存のＰＬＢＡ，Ｓｔｒｉｐ　Ｓｉｚｅ、Ｓｔｒｉｐ　Ｄｅｐｔｈ、Ｓｔｒｉｐｅ　Ｓｉｚｅ，メンバーデイスク数を用いたＲＡＩＤマッピング処理を利用して、ＬＤＥ処理を実行できる。又、増設処理とＲＡＩＤレベル変換処理を同じ処理で実現できる。

【０１１９】

[容量増加処理／ＲＡＩＤレベル変換処理]

図２４は、容量増加時の旧ＲＡＩＤ定義と新ＲＡＩＤ定義の説明図、図２５は、その動作説明図である。図２４に示すように、旧（テンポラリ）ＲＡＩＤ定義８０では、ＲＡＩＤレベル５で、３＋１（ＰＬＵＮｓ１０－１３）で、メンバーでイスク数４、ブロック数４００を定義する。一方、新ＲＡＩＤ定義８２では、ＲＡＩＤレベル５で、４＋１（ＰＬＵＮｓ１０－１４）で、メンバーデイスク数

５、ブロック数５００を定義する。

【０１２０】

前述の図４及び図５の説明のように、メンバーデイスク数の増加で、マッピングが変更され、図２５に示すように、ＲＡＩＤ５（３＋１）が、ＲＡＩＤ５（４＋１）に増設される。

【０１２１】

図２６は、ＲＡＩＤレベル変換時の旧ＲＡＩＤ定義と新ＲＡＩＤ定義の説明図、図２７は、その動作説明図、図２８は、ＲＬＵテーブルとＤＬＵテーブルの遷移の説明図である。

【０１２２】

図２６に示すように、旧（テンポラリ）ＲＡＩＤ定義８０では、ＲＡＩＤレベル０＋１で、メンバーデイスク数４（ＰＬＵＮｓ０－３）で、ブロック数４００を定義する。一方、新ＲＡＩＤ定義８２では、ＲＡＩＤレベル５で、３＋１（ＰＬＵＮｓ０－３）で、メンバーデイスク数４、ブロック数４００を定義する。

【０１２３】

前述の図６及び図７の説明のように、ＲＡＩＤ０＋１のストライプをリードし、図４及び図５の説明のように、ＲＡＩＤ５の定義で、マッピングが変更され、図２７に示すように、ＲＡＩＤ０＋１（４＋４）が、ＲＡＩＤ５（３＋１）にレベル変換される。これにより、図２８に示すように、ＲＬＵテーブル７２、ＤＬＵテーブル７４が、ＲＡＩＤ０＋１からＲＡＩＤ５の定義に変更される。

【０１２４】

[他の実施の形態]

図２９は、本発明の他のアドレス変換の説明図である。この実施の形態では、予め、変換の対象となるＲＡＩＤ構成定義間のＲＬＢＡの差分をテーブルに求めておき、旧ＲＡＩＤ定義のＲＬＢＡで、旧ＲＡＩＤ定義のリードＰＬＢＡを計算する。

【０１２５】

この旧ＲＡＩＤ定義のＲＬＢＡを、ＲＬＢＡ間の差分を足し、新ＲＬＢＡを求め、新ＲＡＩＤ定義８２で、ライトＰＬＢＡを計算する。このようにすると、Ｒ

ＬＢＡ間の差分計算の手間はかかるが、高速にマッピング変換が可能となる。

【０１２６】

前述の実施の形態では、２ＣＭ（コントローラ）－４ＤＥ（デバイスエンクロージャ）の構成で説明したが、４台のコントローラを有する４ＣＭ-１６ＤＥの構成でも、各ＣＭは、保守ＰＣ３のＬＤＥ起動により、同様に、ＬＤＥ処理できる。

【０１２７】

又、前述の実施の形態では、図１２のようなＲＡＩＤレベルで説明したが、これ以外のＲＡＩＤレベル（ＲＡＩＤ２，ＲＡＩＤ３，ＲＡＩＤ４）のストレージシステムに適用できる。又、物理デイスクは、磁気デイスク、光デイスク、光磁気デイスク、各種のストレージデバイスを適用できる。

【０１２８】

以上、本発明を実施の形態により説明したが、本発明の趣旨の範囲内において、本発明は、種々の変形が可能であり、本発明の範囲からこれらを排除するものではない。

【０１２９】

（付記１）ＲＡＩＤ構成定義に従い、データを分解し、複数の物理デイスク装置に、並列にリード／ライトするＲＡＩＤ装置において、上位装置からのＩ／Ｏ要求に応じて、前記ＲＡＩＤ構成定義によるＲＬＵマッピングに従い、前記複数の物理デイスク装置をアクセスする制御部と、少なくとも旧ＲＡＩＤレベルと旧論理デバイス数を定義した旧ＲＡＩＤ構成定義情報と、少なくとも新ＲＡＩＤレベルと新論理デバイス数を定義した新ＲＡＩＤ構成定義情報とを格納するテーブルと、旧ＲＡＩＤ構成から新ＲＡＩＤ構成に変更するため、データを一時格納するためのキャッシュメモリとを有し、前記制御部は、前記テーブルの前記旧ＲＡＩＤ構成定義によるＲＬＵマッピングに従い、前記複数の物理デイスク装置からデータを前記キャッシュメモリに読出し、前記テーブルの前記新ＲＡＩＤ構成定義によるＲＬＵマッピングに従い、前記キャッシュメモリに読み出したデータを、前記複数の物理デイスク装置にライトすることを特徴とするＲＡＩＤ装置。

【０１３０】

（付記２）前記制御部は、前記旧ＲＡＩＤ構成定義のＲＡＩＤレベルによるＲＬＵマッピングに従い、前記複数の物理デイスク装置からデータを前記キャッシュメモリに読出し、前記新ＲＡＩＤ構成定義のＲＡＩＤレベルによるＲＬＵマッピングに従い、前記キャッシュメモリに読み出したデータを、前記複数の物理デイスク装置にライトするＲＡＩＤレベル変換処理を行うことを特徴とする付記１のＲＡＩＤ装置。

【０１３１】

（付記３）前記制御部は、前記旧ＲＡＩＤ構成定義の前記論理デバイス数によるＲＬＵマッピングに従い、前記複数の物理デイスク装置からデータを前記キャッシュメモリに読出し、前記新ＲＡＩＤ構成定義の前記論理デバイス数によるＲＬＵマッピングに従い、前記キャッシュメモリに読み出したデータを、前記複数の物理デイスク装置にライトする容量増加処理を行うことを特徴とする付記１のＲＡＩＤ装置。

【０１３２】

（付記４）前記制御部は、前記旧ＲＡＩＤ構成から前記新ＲＡＩＤ構成への変換をシーケンシャルに実行し、且つその進捗状況を管理するとともに、前記変換中に、前記上位装置からのＩ／Ｏ要求に対し、変換済み領域かを判定し、変換済み領域に対しては、前記新ＲＡＩＤ構成定義で、前記Ｉ／Ｏ要求を実行し、変換済みでない領域に対しては、前記旧ＲＡＩＤ構成定義で、前記Ｉ／Ｏ要求を実行することを特徴とする付記１のＲＡＩＤ装置。

【０１３３】

（付記５）前記制御部は、新ＲＡＩＤ構成定義によるＲＬＢＡを上位のＬＢＡに変換した後、前記上位のＬＢＡから前記旧ＲＡＩＤ構成定義によるＲＬＵマッピングに従い、前記複数の物理デイスク装置からデータを前記キャッシュメモリに読出し、前記ＲＬＢＡから前記新ＲＡＩＤ構成定義によるＲＬＵマッピングに従い、前記キャッシュメモリに読み出したデータを、前記複数の物理デイスク装置にライトすることを特徴とする付記１のＲＡＩＤ装置。

【０１３４】

（付記６）前記制御部は、前記旧ＲＡＩＤ構成から前記新ＲＡＩＤ構成への変

換後、前記旧ＲＡＩＤ構成定義を前記テーブルから削除することを特徴とする付記１のＲＡＩＤ装置。

【０１３５】

（付記７）前記制御部は、指示された新ＲＡＩＤ構成定義のパラメータと前記旧ＲＡＩＤ構成定義とに従い、前記新ＲＡＩＤ構成定義を、前記テーブルに作成することを特徴とする付記１のＲＡＩＤ装置。

【０１３６】

（付記８）前記制御部は、変換領域に対応する前記キャッシュメモリの領域を獲得した後、前記旧ＲＡＩＤ構成から前記新ＲＡＩＤ構成への変換をシーケンシャルに実行することを特徴とする付記１のＲＡＩＤ装置。

【０１３７】

（付記９）前記制御部は、前記変換領域に対応する前記キャッシュメモリの領域が獲得できない時は、前記変換処理を複数回に分けて実行することを特徴とする付記８のＲＡＩＤ装置。

【０１３８】

（付記１０）前記制御部は、前記ＲＡＩＤ構成のストライプに対応したストリップデプス、ストライプサイズに応じて、前記ＲＬＵマッピングすることを特徴とする付記１のＲＡＩＤ装置。

【０１３９】

（付記１１）ＲＡＩＤ構成定義に従い、データを分解し、複数の物理デイスク装置に、並列にリード／ライトするＲＡＩＤ装置の論理デバイス拡張方法において、少なくとも旧ＲＡＩＤレベルと旧論理デバイス数を定義した旧ＲＡＩＤ構成定義情報によるＲＬＵマッピングに従い、前記複数の物理デイスク装置からデータを前記キャッシュメモリに読出すステップと、少なくとも新ＲＡＩＤレベルと新論理デバイス数を定義した新ＲＡＩＤ構成定義情報によるＲＬＵマッピングに従い、前記キャッシュメモリに読み出したデータを、前記複数の物理デイスク装置にライトすることを特徴とする論理デバイス拡張方法。

【０１４０】

（付記１２）前記読出しステップは、前記旧ＲＡＩＤ構成定義のＲＡＩＤレベ

ルによるＲＬＵマッピングに従い、前記複数の物理デイスク装置からデータを前記キャッシュメモリに読出すステップからなり、前記ライトステップは、前記新ＲＡＩＤ構成定義のＲＡＩＤレベルによるＲＬＵマッピングに従い、前記キャッシュメモリに読み出したデータを、前記複数の物理デイスク装置にライトするステップからなることを特徴とする付記１１の論理デバイス拡張方法。

【０１４１】

（付記１３）前記読出しステップは、前記旧ＲＡＩＤ構成定義の前記論理デバイス数によるＲＬＵマッピングに従い、前記複数の物理デイスク装置からデータを前記キャッシュメモリに読出すステップからなり、前記ライトステップは、前記新ＲＡＩＤ構成定義の前記論理デバイス数によるＲＬＵマッピングに従い、前記キャッシュメモリに読み出したデータを、前記複数の物理デイスク装置にライトする容量増加ステップからなることを特徴とする付記１１の論理デバイス拡張方法。

【０１４２】

（付記１４）前記旧ＲＡＩＤ構成から前記新ＲＡＩＤ構成への変換をシーケンシャルに実行するその進捗状況を管理するステップと、前記変換中に、前記上位装置からのＩ／Ｏ要求に対し、変換済み領域かを判定するステップと、変換済み領域に対しては、前記新ＲＡＩＤ構成定義で、前記Ｉ／Ｏ要求を実行するステップと、変換済みでない領域に対しては、前記旧ＲＡＩＤ構成定義で、前記Ｉ／Ｏ要求を実行するステップとからなることを特徴とする付記１１の論理デバイス拡張方法。

【０１４３】

（付記１５）前記読出しステップは、新ＲＡＩＤ構成定義によるＲＬＢＡを上位のＬＢＡに変換した後、前記上位のＬＢＡから前記旧ＲＡＩＤ構成定義によるＲＬＵマッピングに従い、前記複数の物理デイスク装置からデータを前記キャッシュメモリに読出すステップからなり、前記ライトステップは、前記ＲＬＢＡから前記新ＲＡＩＤ構成定義によるＲＬＵマッピングに従い、前記キャッシュメモリに読み出したデータを、前記複数の物理デイスク装置にライトするステップからなることを特徴とする付記１１の論理デバイス拡張方法。

【０１４４】

（付記１６）前記旧ＲＡＩＤ構成から前記新ＲＡＩＤ構成への変換後、前記旧ＲＡＩＤ構成定義を削除するステップを更に有することを特徴とする付記１１の論理デバイス拡張方法。

【０１４５】

（付記１７）指示された新ＲＡＩＤ構成定義のパラメータと前記旧ＲＡＩＤ構成定義とに従い、前記新ＲＡＩＤ構成定義を、テーブルに作成するステップを更に有することを特徴とする付記１１の論理デバイス拡張方法。

【０１４６】

（付記１８）変換領域に対応する前記キャッシュメモリの領域を獲得した後、前記旧ＲＡＩＤ構成から前記新ＲＡＩＤ構成への変換をシーケンシャルに実行するステップを更に有することを特徴とする付記１１の論理デバイス拡張方法。

【０１４７】

（付記１９）前記変換領域に対応する前記キャッシュメモリの領域が獲得できない時は、前記変換処理を複数回に分けて実行するステップを更に有することを特徴とする付記１８の論理デバイス拡張方法。

【０１４８】

（付記２０）前記ＲＡＩＤ構成のストライプに対応したストリップデプス、ストライプサイズに応じて、前記ＲＬＵマッピングするステップを更に有することを特徴とする付記１１の論理デバイス拡張方法。

【０１４９】

【発明の効果】

このように、本発明では、少なくとも、ＲＡＩＤレベルと論理デバイス数を定義した新旧のＲＡＩＤ構成定義情報を使用し、それぞれによりＲＬＵマッピングして、ＲＡＩＤ構成を変更するため、多様なＲＡＩＤレベルの変換、容量増加を活性状態で実現できる。

【図面の簡単な説明】

【図１】

本発明の一実施の形態のストレージシステムの構成図である。

【図２】

図１の制御モジュールの構成図である。

【図３】

図２のＲＡＩＤ論理空間と論理テーブルの説明図である。

【図４】

図２のＲＡＩＤ５の論理マッピング図である。

【図５】

図４のＲＡＩＤ５のＰＬＢＡ計算処理フロー図である。

【図６】

図２のＲＡＩＤ０＋１の論理マッピング図である。

【図７】

図６のＲＡＩＤ０＋１のＰＬＢＡ計算処理フロー図である。

【図８】

図１のＲＡＩＤグループ容量増加の一実施の形態の説明図である。

【図９】

図１のＲＡＩＤグループ容量増加の他の実施の形態の説明図である。

【図１０】

図１のＲＡＩＤレベル変換の一実施の形態の説明図である。

【図１１】

図１のＲＡＩＤレベル変換の他の実施の形態の説明図である。

【図１２】

図１のＲＡＩＤレベル変換の関係図である。

【図１３】

図２のＬＤＥ全体処理フロー図である。

【図１４】

図１３のＲＡＩＤ構成定義の説明図である。

【図１５】

図２のＬＤＥ制御モジュールの構成図である。

【図１６】

図１５のＬＤＥマッピング処理の説明図である。

【図１７】

図１５のＬＤＥ詳細処理フロー図である。

【図１８】

図１７のキャッシュメモリ獲得処理フロー図である。

【図１９】

図１７のシーケンシャルＬＤＥの説明図である。

【図２０】

図１５のＩ／Ｏ競合処理フロー図である。

【図２１】

図１５のリードアドレス計算処理フロー図である。

【図２２】

図１５のライトアドレス計算処理フロー図である。

【図２３】

図２１及び図２２のアドレス変換処理の説明図である。

【図２４】

図８の容量増加の新旧ＲＡＩＤ定義の説明図である。

【図２５】

図２４の新旧ＲＡＩＤ定義による容量増加の説明図である。

【図２６】

図１０のＲＡＩＤレベル変換の新旧ＲＡＩＤ定義の説明図である。

【図２７】

図２６の新旧ＲＡＩＤ定義によるＲＡＩＤレベル変換の説明図である。

【図２８】

図２７のＲＡＩＤレベル変換によるＲＬＵテーブル、ＤＬＵテーブルの遷移の説明図である。

【図２９】

本発明の他の実施の形態のリード／ライトアドレス変換の説明図である。

【符号の説明】

１、２　ストレージコントローラ

３　保守ＰＣ

１１、１２、２１、２３　チャネルアダプター

１３、１４、２３、２４　デバイスアダプター

１０、２０　ＣＰＵ

１５，２５　メモリ

１６、２６　キャッシュメモリ

３０、４０　ＰＣＩ－ノードブリッジ回路

３１、４１　ＰＣＩバス

５０、５０－１～５０－ｍ、５２－１～５２－ｎ　物理デイスク装置（ストレージ装置）

３４　保守エージェント

３５　リソーススレッド

３７　キャッシュスレッド

３８　ＯＶＳＭスレッド

３９　ＲＡＩＤスレッド

４８　システム制御スレッド

７０　ＯＬＵテーブル

７２　ＲＬＵテーブル

７４　ＤＬＵテーブル

７６　ＰＬＵテーブル

８０　テンポラリＲＡＩＤ定義

８２　ＲＡＩＤ定義

９０　キャッシュメモリ獲得処理

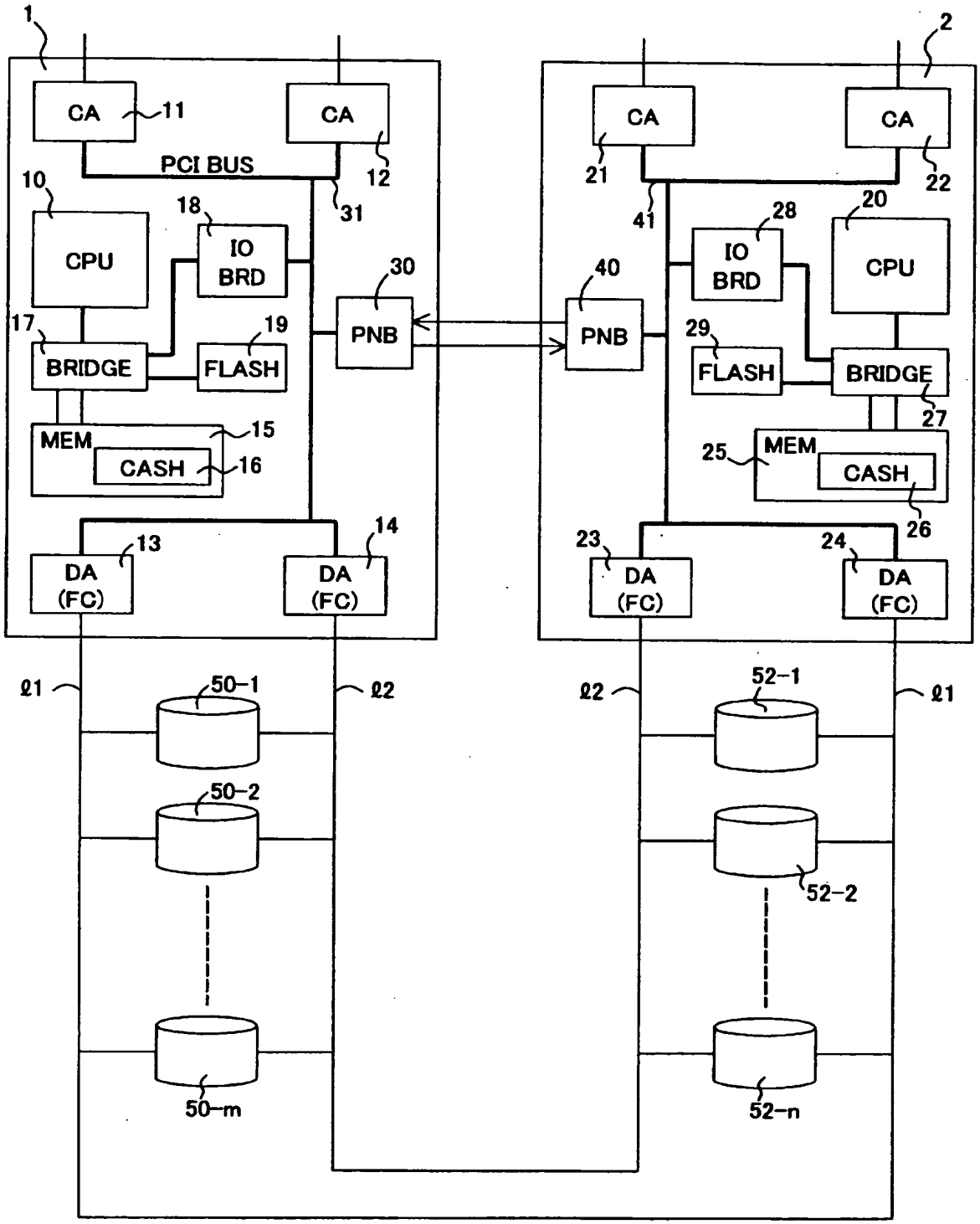
９２　進捗状況管理処理

９４　リード処理

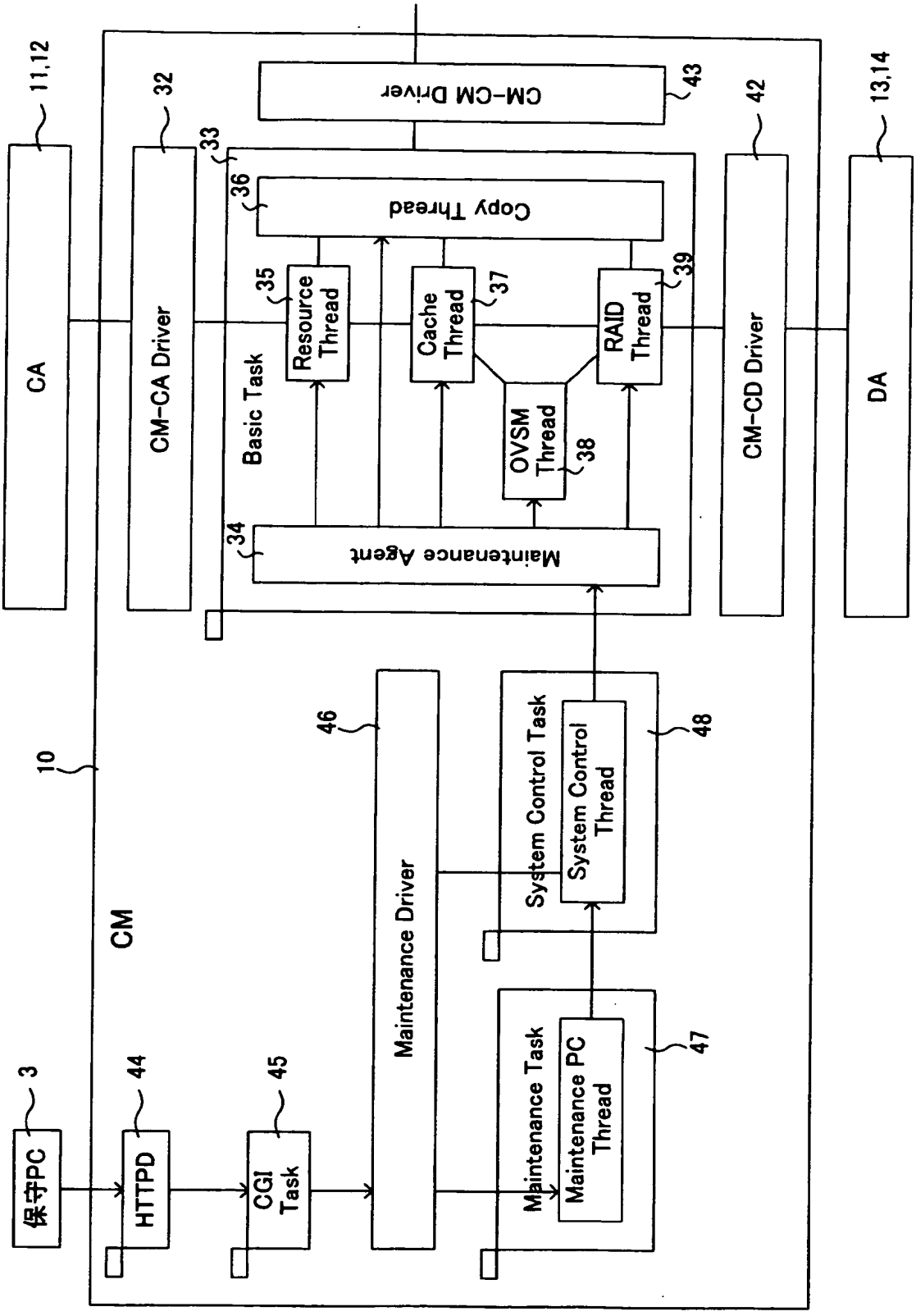
９６　ライト処理

【書類名】　　図面

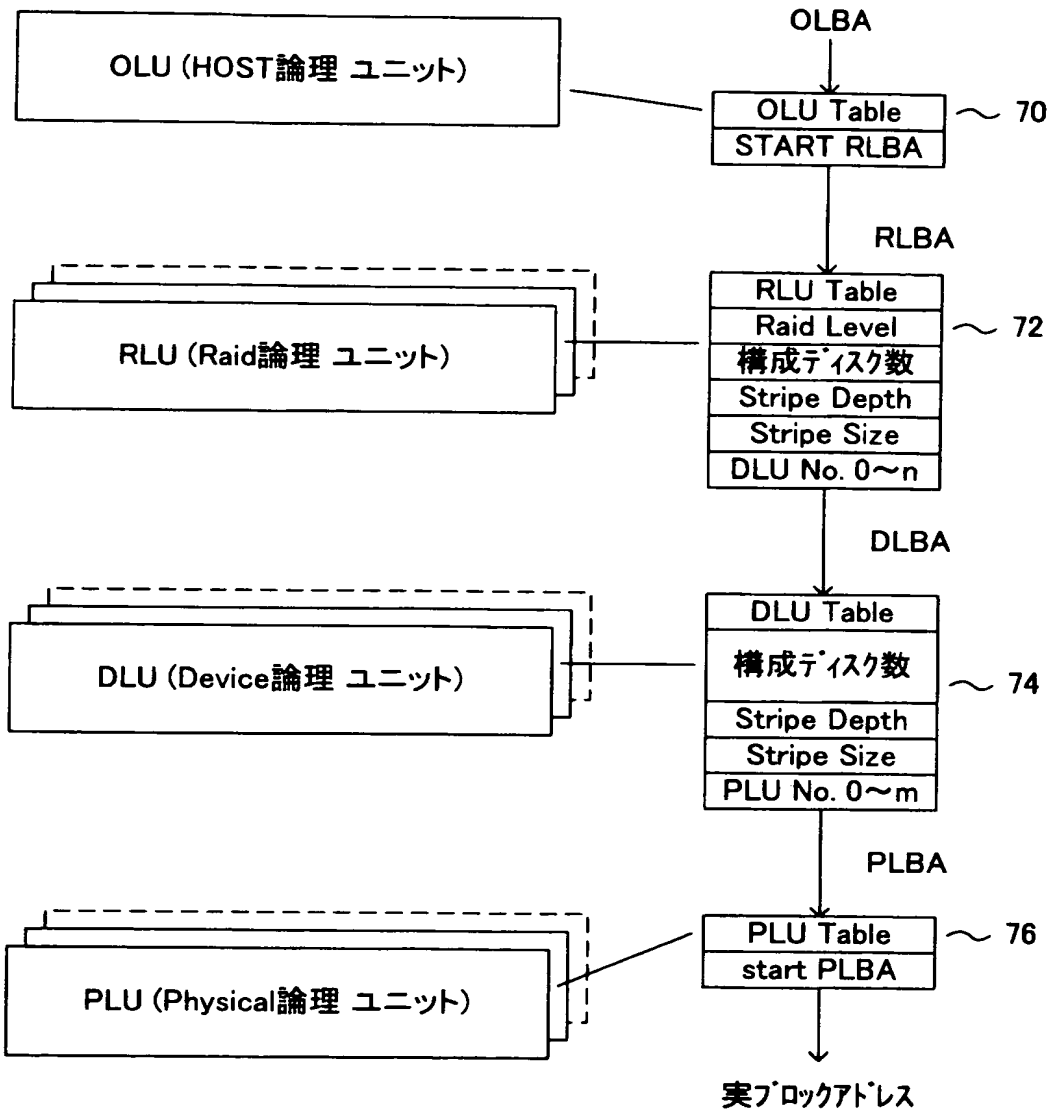
【図１】


1
2
CA
11
CA
12
PCI BUS
31
CA
21
CA
22
41
10
CPU
18
IO BRD
30
PNB
40
PNB
28
IO BRD
20
CPU
17
BRIDGE
19
FLASH
29
FLASH
BRIDGE
27
MEM
15
CASH
16
25
MEM
CASH
26
13
DA (FC)
14
DA (FC)
23
DA (FC)
24
DA (FC)
ℓ1
50-1
ℓ2
ℓ2
52-1
ℓ1
50-2
52-2
50-m
52-n

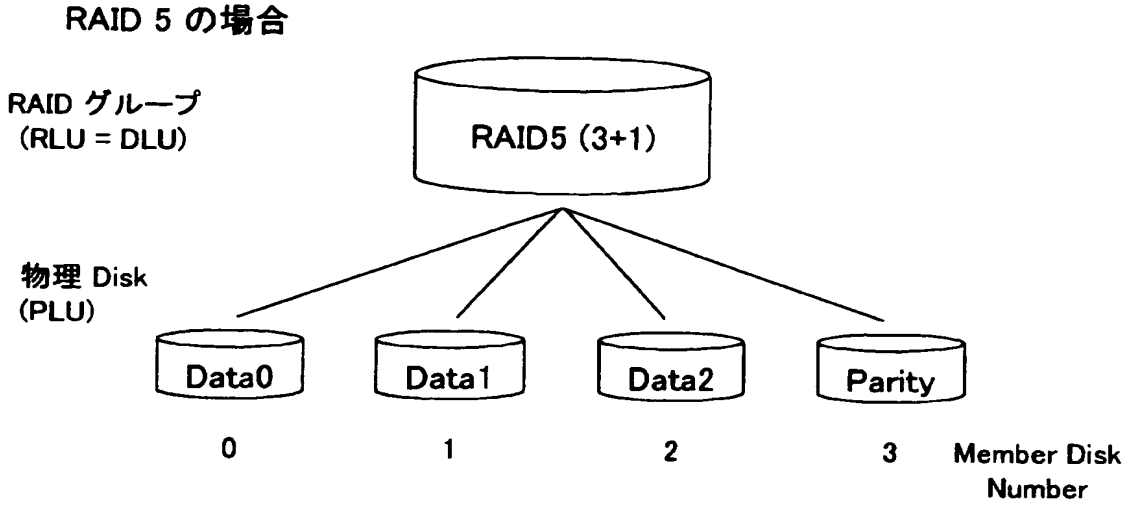
【図２】

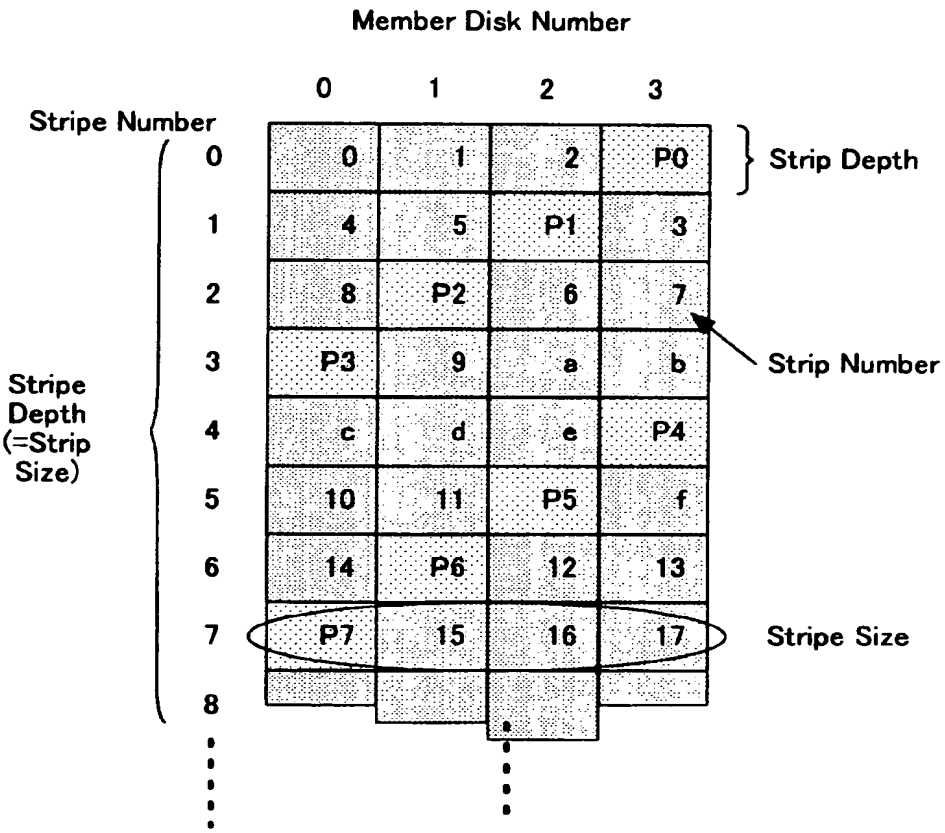

11,12
32
33
43
42
13,14
CM-CM Driver
Copy Thread
36
CA
CM-CA Driver
Basic Task
35
Resource Thread
Cache Thread
37
RAID Thread
39
OVSM Thread
38
CM-CD Driver
DA
Maintenance Agent
34
10
CM
46
Maintenance Driver
System Control Task
System Control Thread
48
3
44
45
保守PC
HTTPD
CGI Task
Maintenance Task
Maintenance PC Thread
47

【図３】

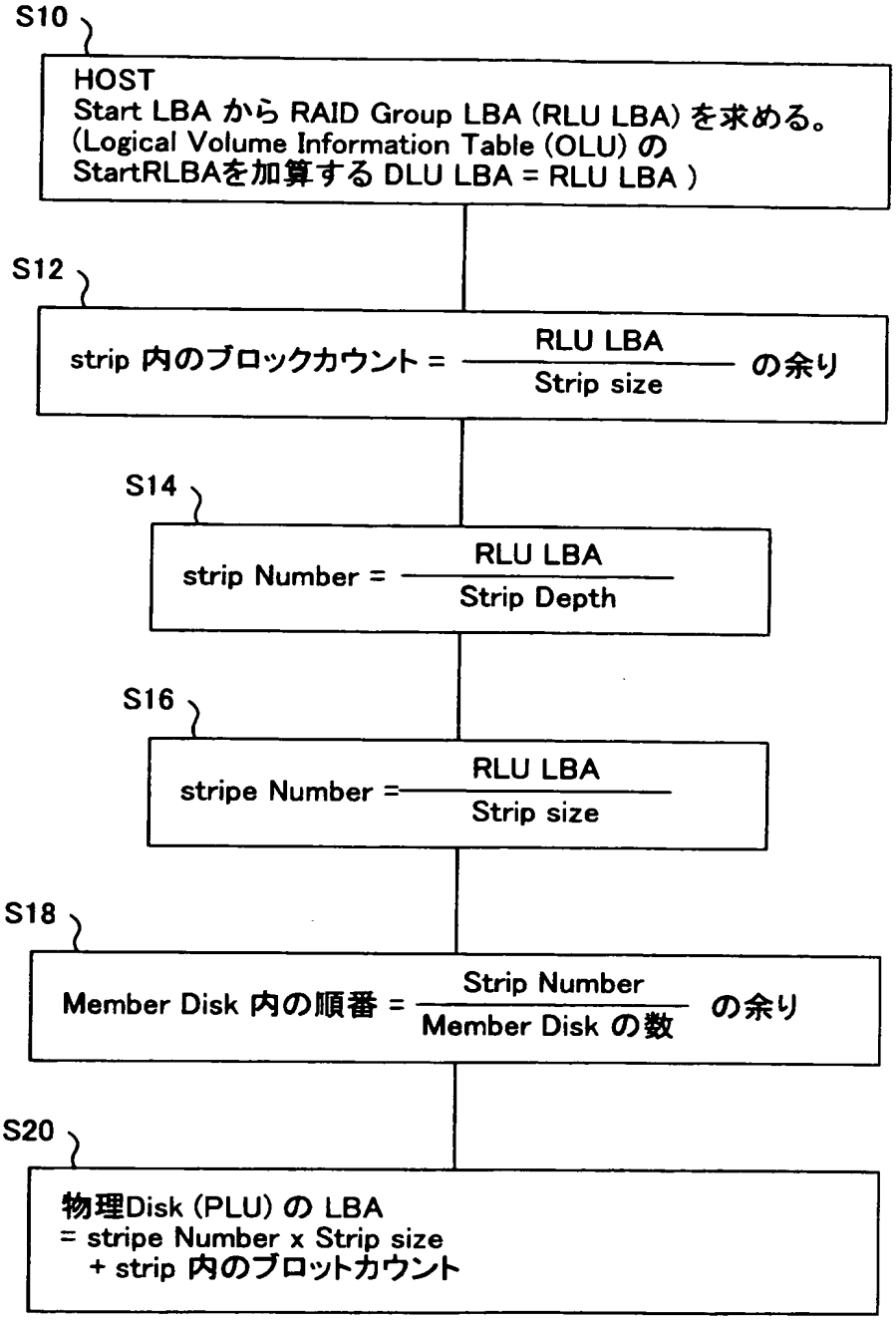

OLU (HOST論理 ユニット)
OLBA
OLU Table
START RLBA
70
RLBA
RLU (Raid論理 ユニット)
RLU Table
Raid Level
構成ディスク数
Stripe Depth
Stripe Size
DLU No. 0～n
72
DLBA
DLU (Device論理 ユニット)
DLU Table
構成ディスク数
Stripe Depth
Stripe Size
PLU No. 0～m
74
PLBA
PLU (Physical論理 ユニット)
PLU Table
start PLBA
76
実ブロックアドレス

【図４】


RAID 5 の場合
RAID グループ
(RLU = DLU)
RAID5 (3+1)
物理 Disk
(PLU)
Data0
Data1
Data2
Parity
0
1
2
3
Member Disk
Number

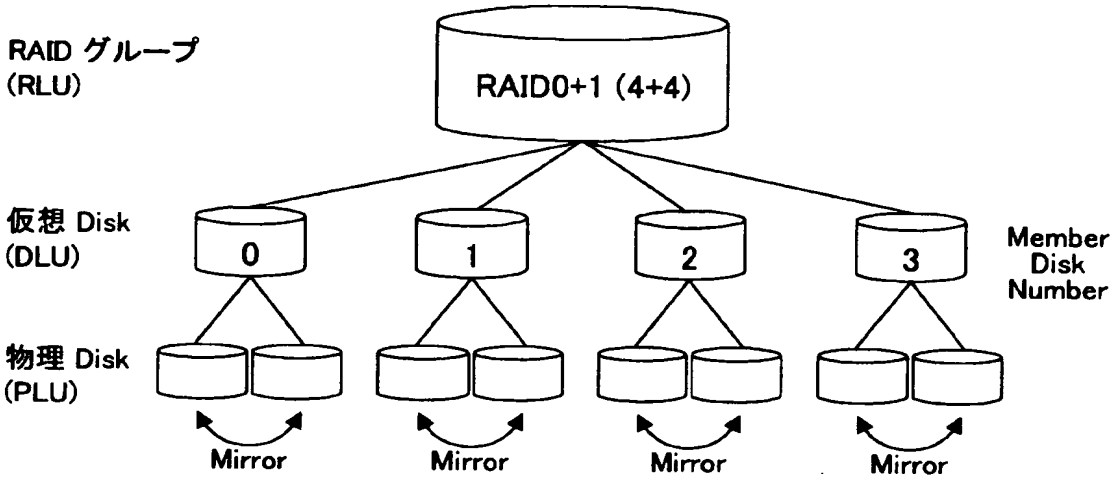

Member Disk Number
0 1 2 3
Stripe Number
0 1 2 P0
4 5 P1 3
8 P2 6 7
P3 9 a b
c d e P4
10 11 P5 f
14 P6 12 13
P7 15 16 17
0
1
2
3
4
5
6
7
8
Strip Depth
Strip Number
Stripe Size
Stripe
Depth
(=Strip
Size)

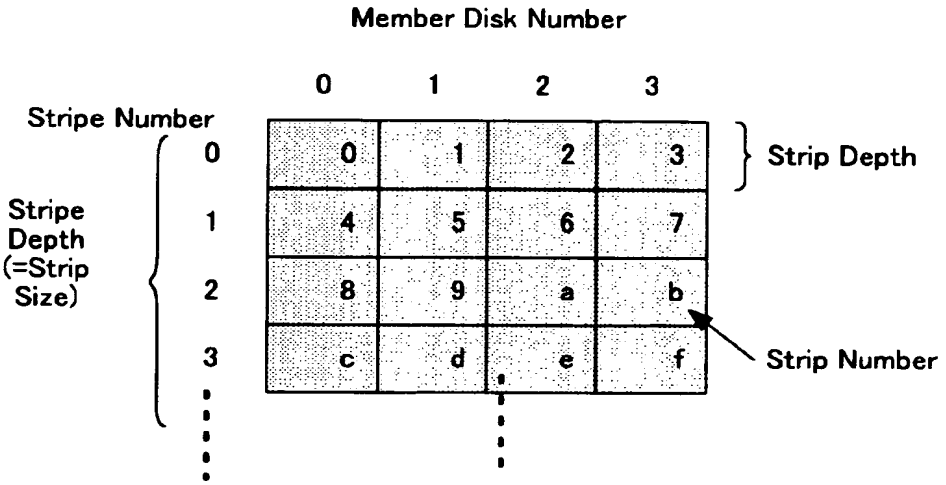
【図５】


S10
HOST
Start LBA から RAID Group LBA (RLU LBA) を求める。
(Logical Volume Information Table (OLU) の
StartRLBAを加算する DLU LBA = RLU LBA)
S12
strip 内のブロックカウント = RLU LBA / Strip size の余り
S14
strip Number = RLU LBA / Strip Depth
S16
stripe Number = RLU LBA / Strip size
S18
Member Disk 内の順番 = Strip Number / Member Disk の数 の余り
S20
物理Disk (PLU) の LBA
= stripe Number x Strip size
+ strip 内のブロットカウント

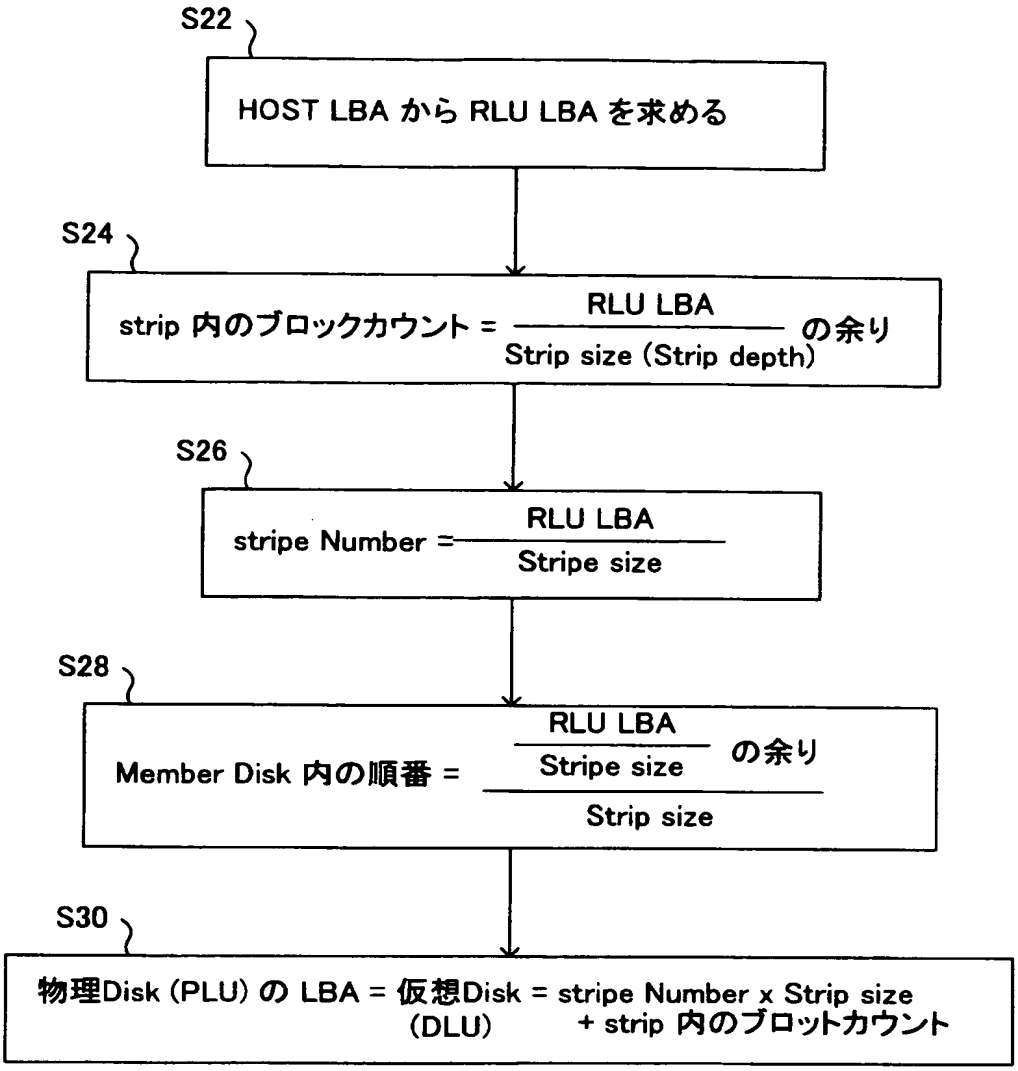
【図６】

RAID 0+1 の場合

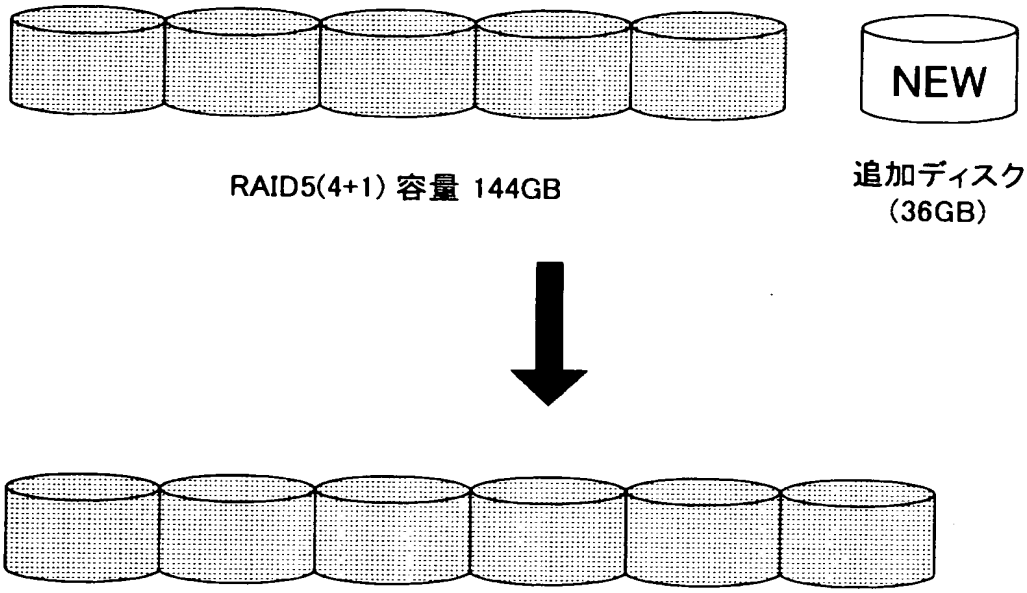




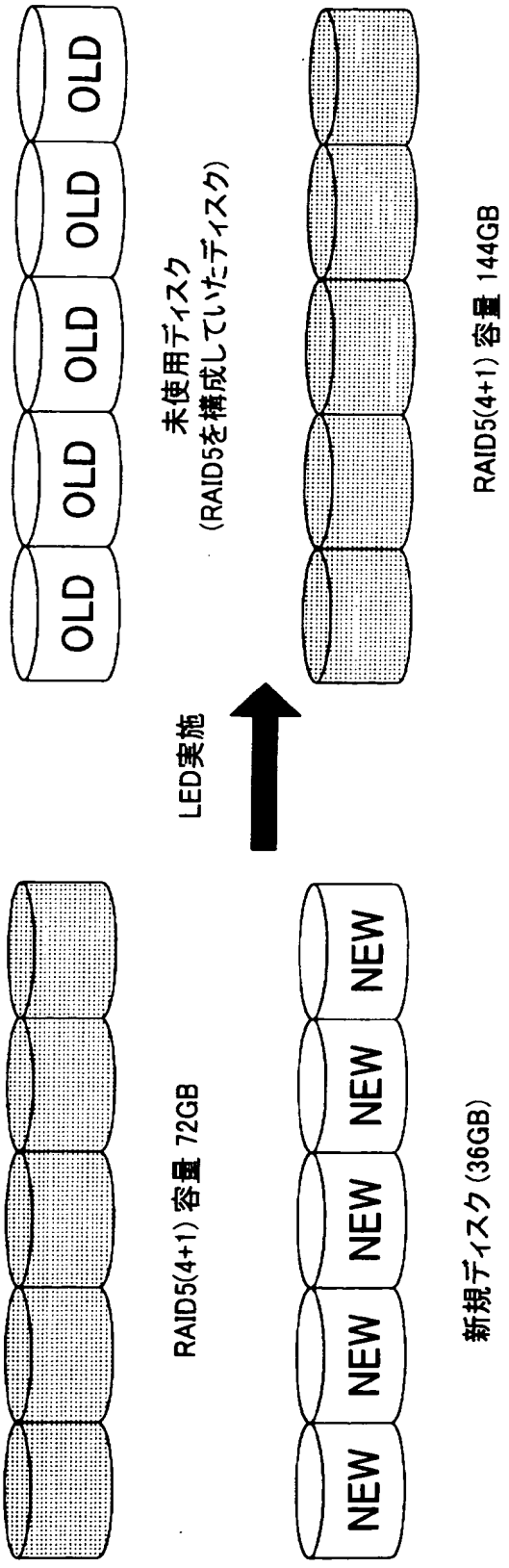
【図７】


S22
HOST LBA から RLU LBA を求める
S24
strip 内のブロックカウント = RLU LBA / Strip size (Strip depth) の余り
S26
stripe Number = RLU LBA / Stripe size
S28
Member Disk 内の順番 = (RLU LBA / Stripe size の余り) / Strip size
S30
物理Disk (PLU) の LBA = 仮想Disk (DLU) = stripe Number x Strip size + strip 内のブロットカウント

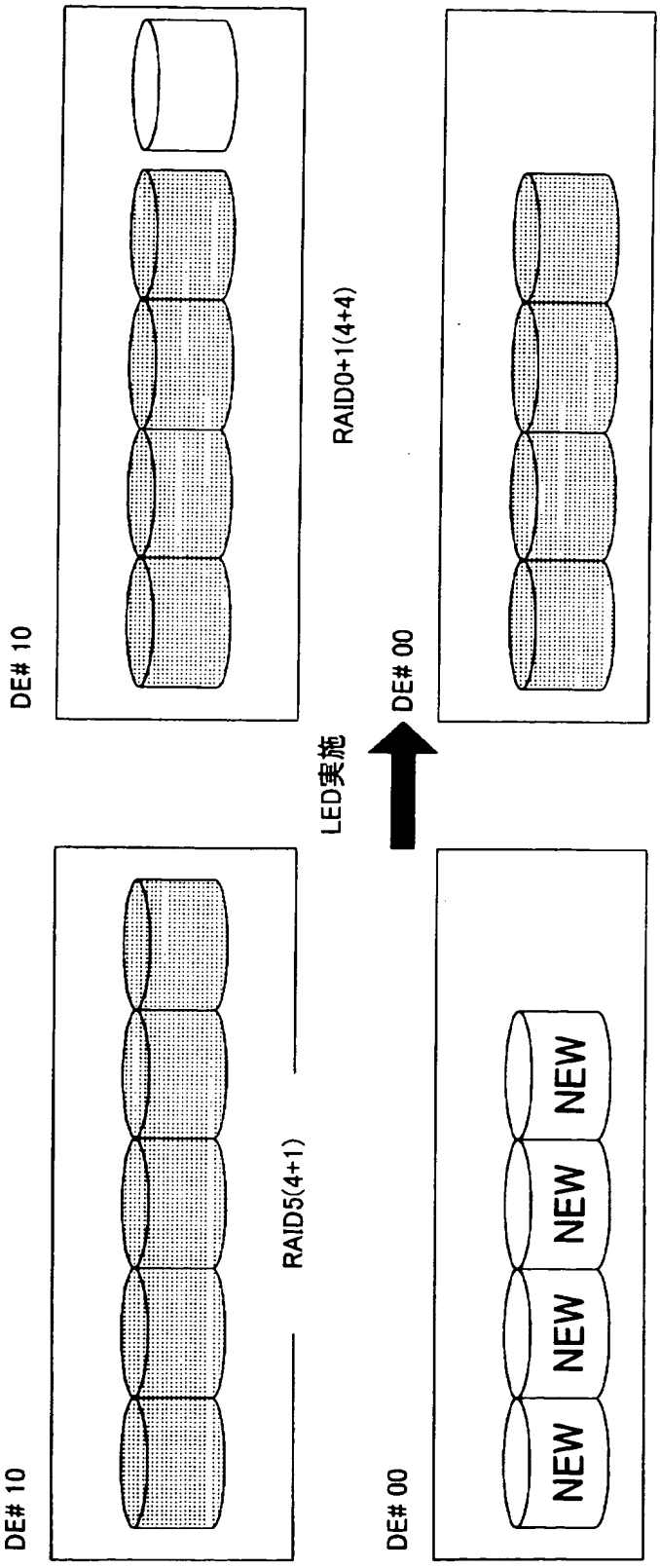
【図８】



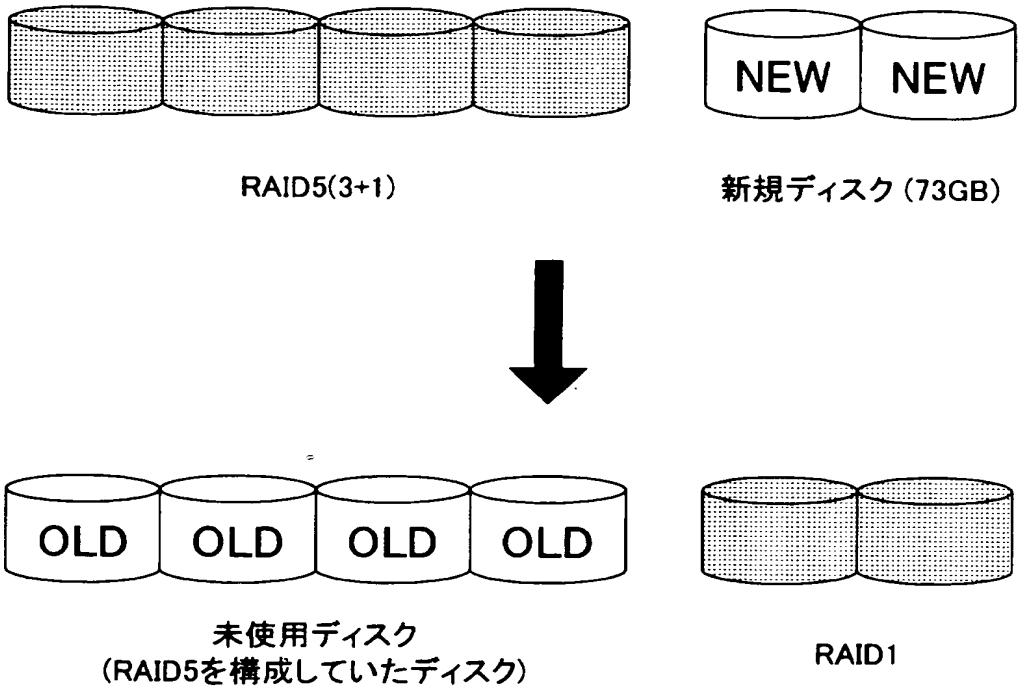
【図９】


OLD
OLD
OLD
OLD
OLD
未使用ディスク
(RAID5を構成していたディスク)
RAID5(4+1) 容量 144GB
LED実施
NEW
NEW
NEW
NEW
NEW
新規ディスク (36GB)
RAID5(4+1) 容量 72GB

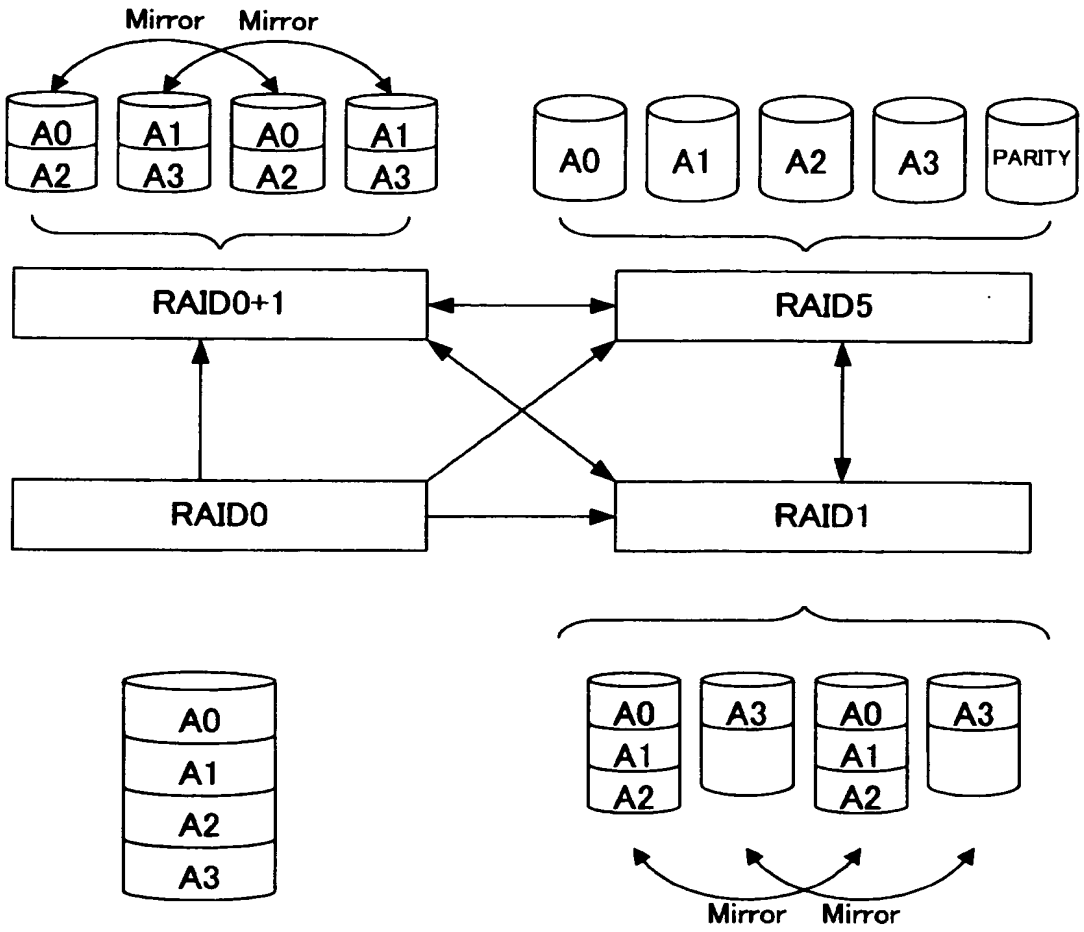
【図１０】



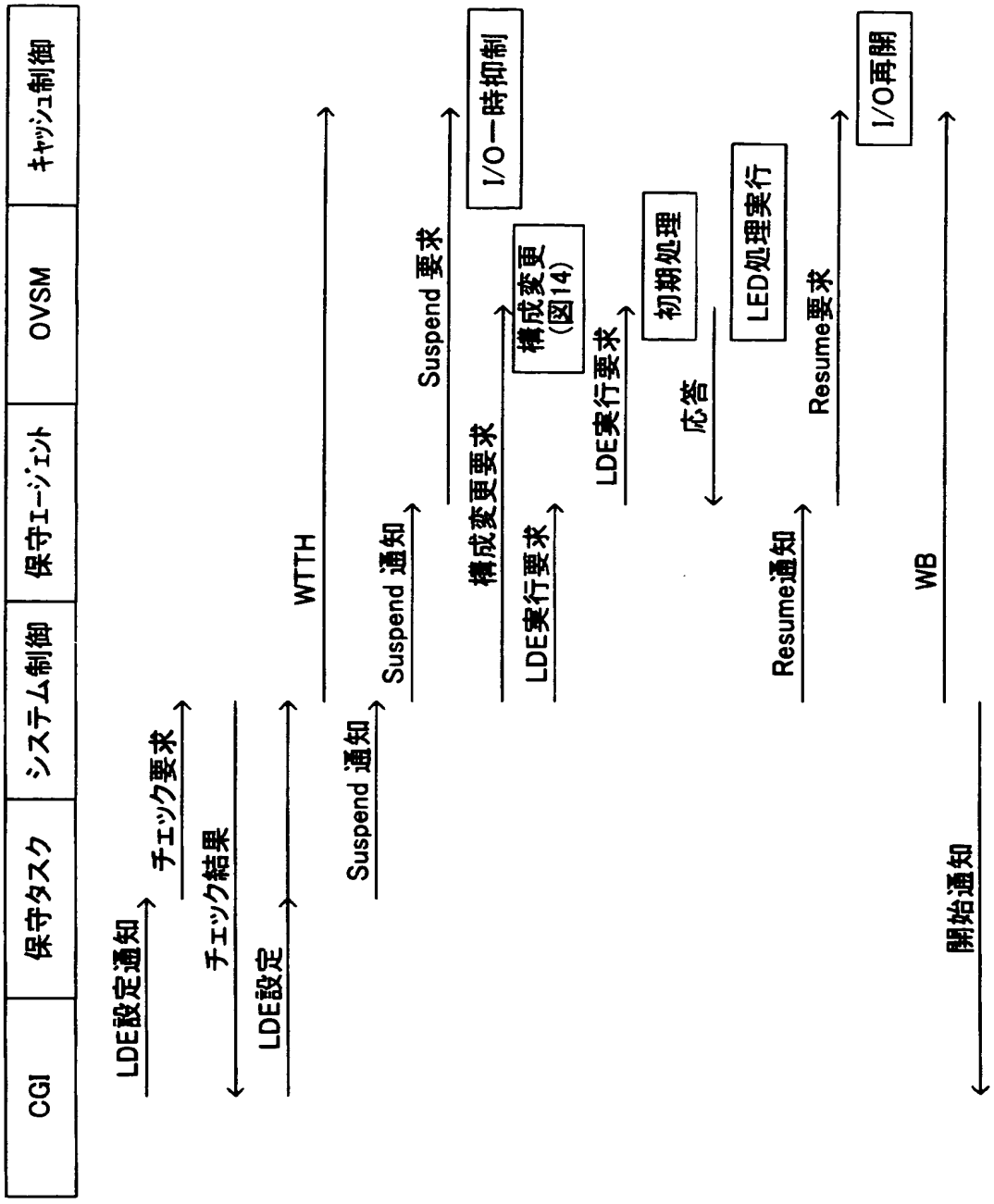
【図１１】


NEW
NEW
RAID5(3+1)
新規ディスク（73GB）
OLD
OLD
OLD
OLD
未使用ディスク
（RAID5を構成していたディスク）
RAID1

【図１２】


Mirror
Mirror
A0
A2
A1
A3
A0
A2
A1
A3
A0
A1
A2
A3
PARITY
RAID0+1
RAID5
RAID0
RAID1
A0
A1
A2
A3
A0
A1
A2
A3
A0
A1
A2
A3
Mirror
Mirror

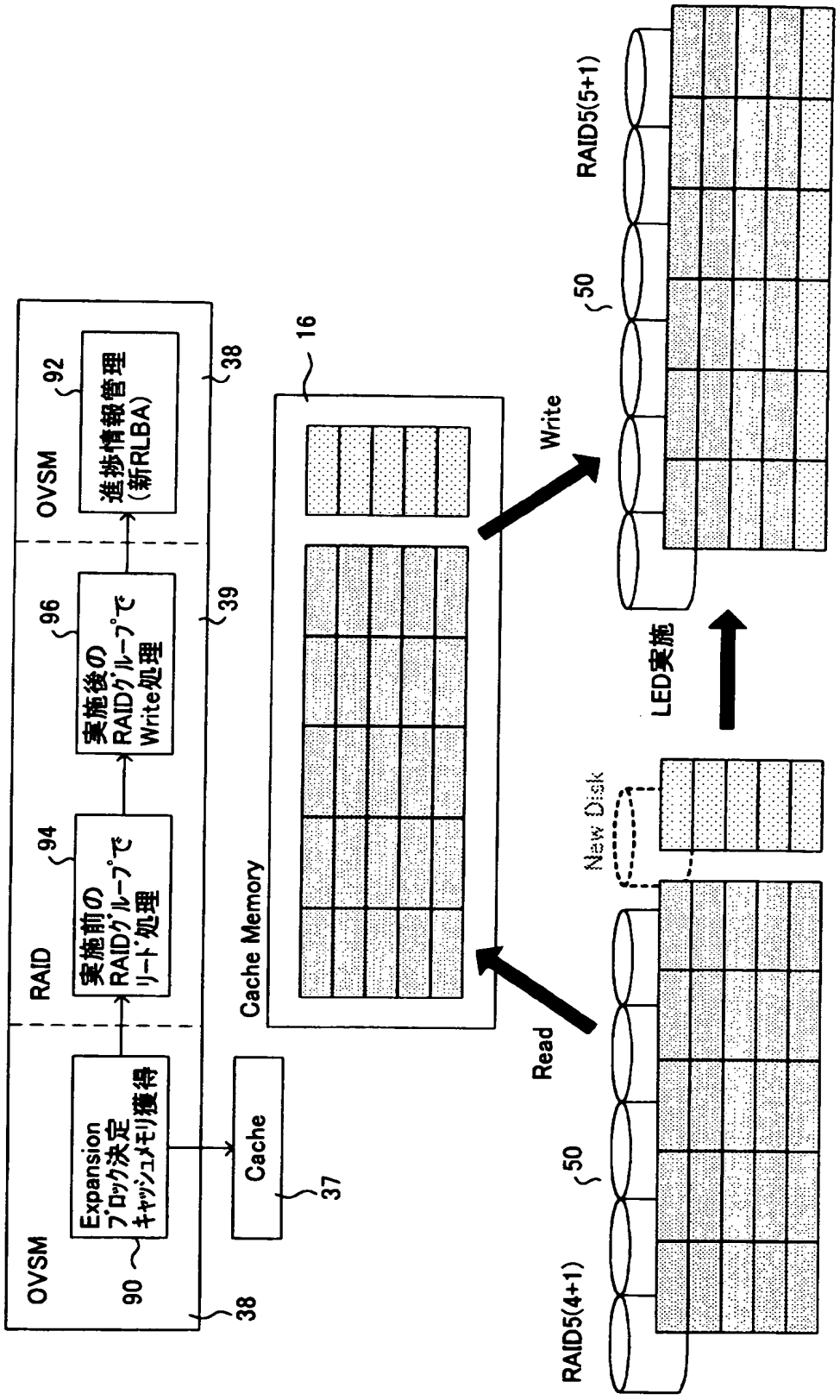
【図１３】


CGI
保守タスク
システム制御
保守エージェント
OVSM
キャッシュ制御
LDE設定通知
チェック要求
チェック結果
LDE設定
WTTH
Suspend 通知
Suspend 通知
Suspend 要求
構成変更要求
LDE実行要求
構成変更
(図14)
I/O一時抑制
LDE実行要求
初期処理
応答
LED処理実行
Resume通知
Resume要求
I/O再開
WB
開始通知

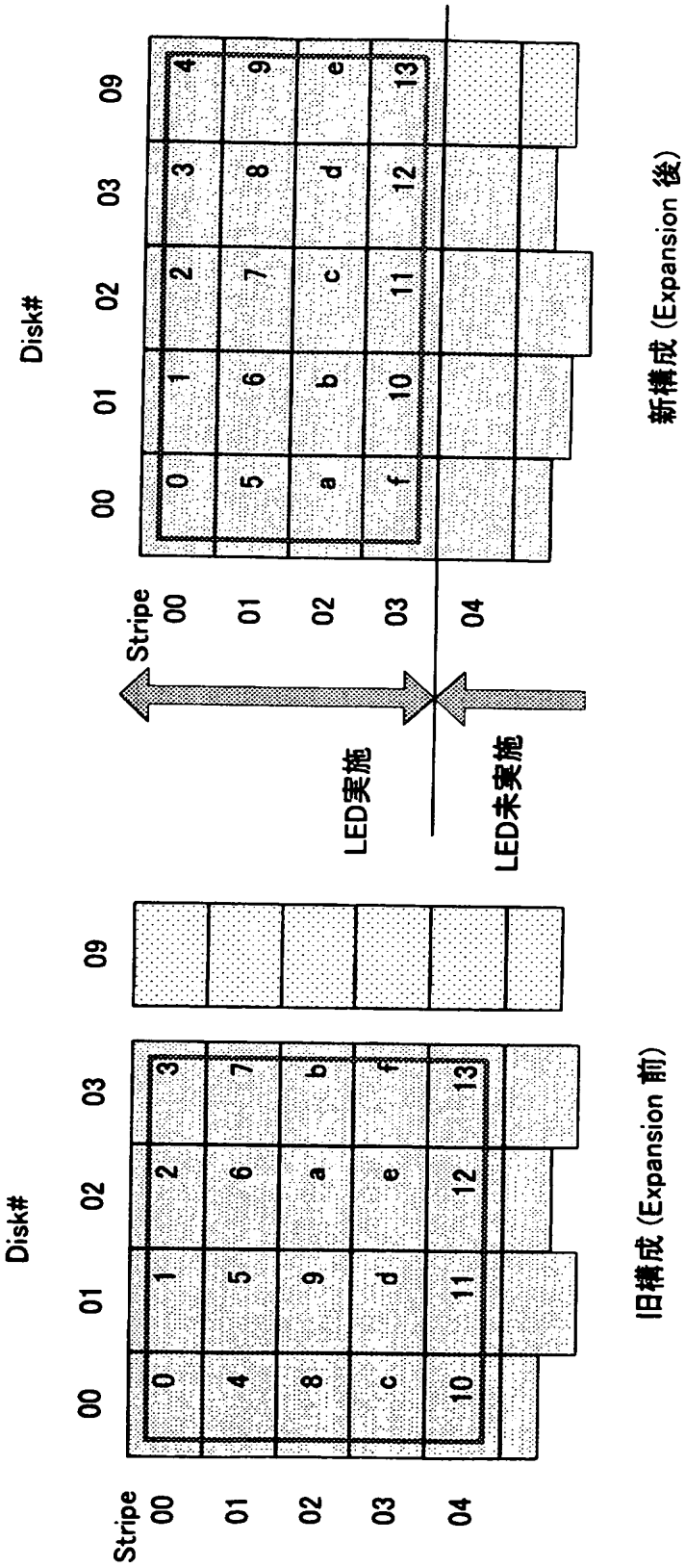
【図１４】


構成定義情報
LUN
項目
RLU
Level
MemberDisks
PLUNs
StartLBA
BlockCount
CheckCode
CVM mode
DLU
DLUN
PLU
-
RAID 定義 (新構成)
82
テンポラリ RAID 定義 (旧構成)
80
LDE状況
実施前
処理中
終了
RAID5
6
10-15
500
LDE Flag
0
RAID5
5
10-14
400
id#0
T Flag
255

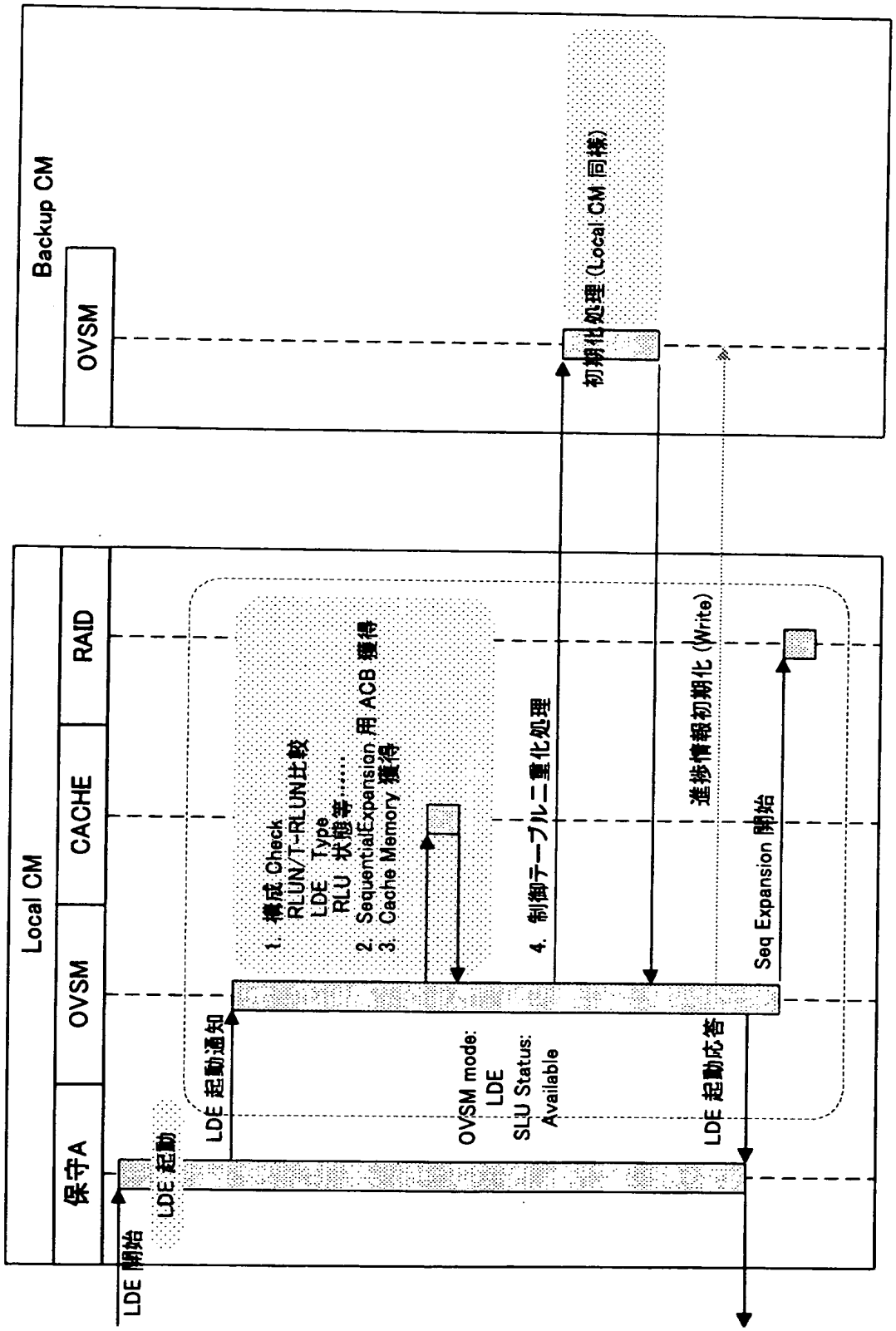
【図１５】


OVSM
RAID
OVSM
38
39
38
90
Expansion
ブロック決定
キャッシュメモリ獲得
94
実施前の
RAIDグループで
リード処理
96
実施後の
RAIDグループで
Write処理
92
進捗情報管理
(新RLBA)
Cache
37
16
Cache Memory
Read
Write
New Disk
LED実施
50
RAID5(4+1)
50
RAID5(5+1)

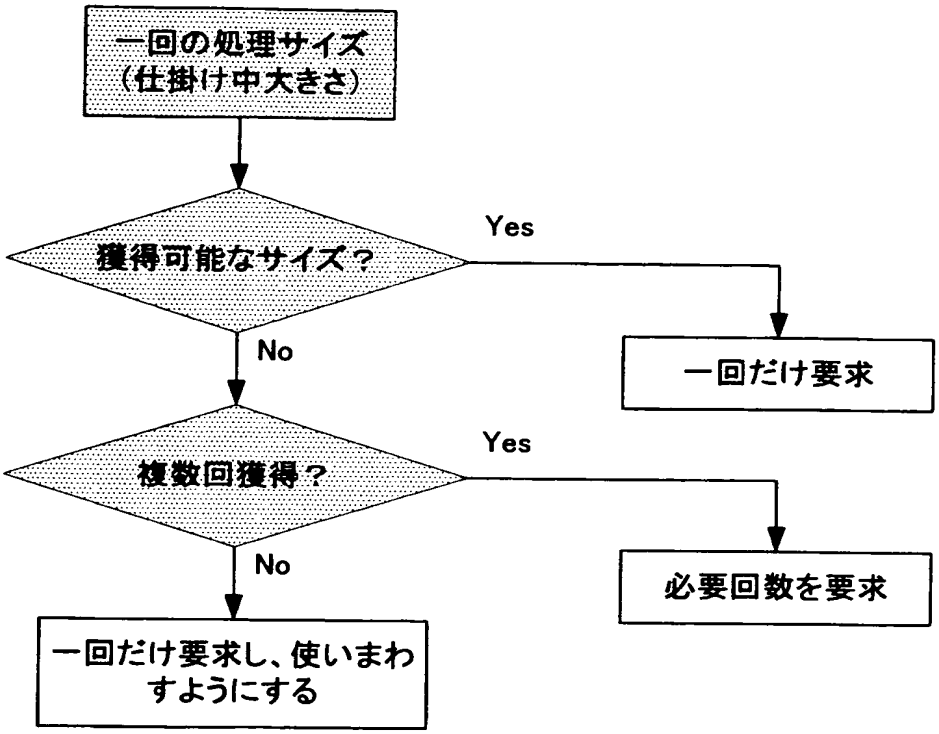
【図１６】


Disk#
00
01
02
03
09
Stripe
00
01
02
03
04
新構成 (Expansion 後)
LED実施
LED未実施
旧構成 (Expansion 前)

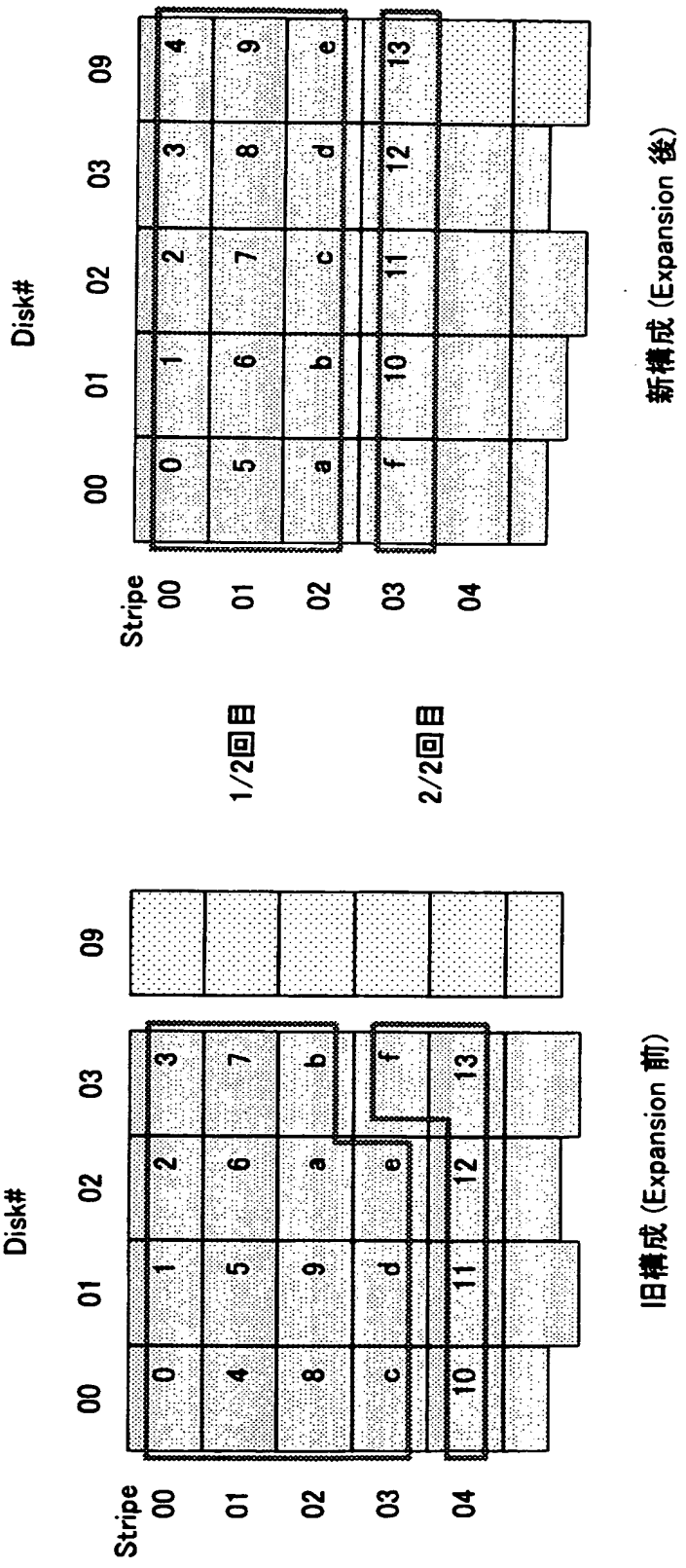
【図１７】


Backup CM
OVSM
初期化処理 (Local CM 同様)
Local CM
保守A
OVSM
CACHE
RAID
LDE 開始
LDE 起動
LDE 起動通知
1. 構成 Check
RLUN/T-RLUN比較
LDE Type
RLU 状態等……
2. SequentialExpansion 用 ACB 獲得
3. Cache Memory 獲得
4. 制御テーブルニ重化処理
OVSM mode:
LDE
SLU Status:
Available
進捗情報初期化 (Write)
Seq Expansion 開始
LDE 起動応答

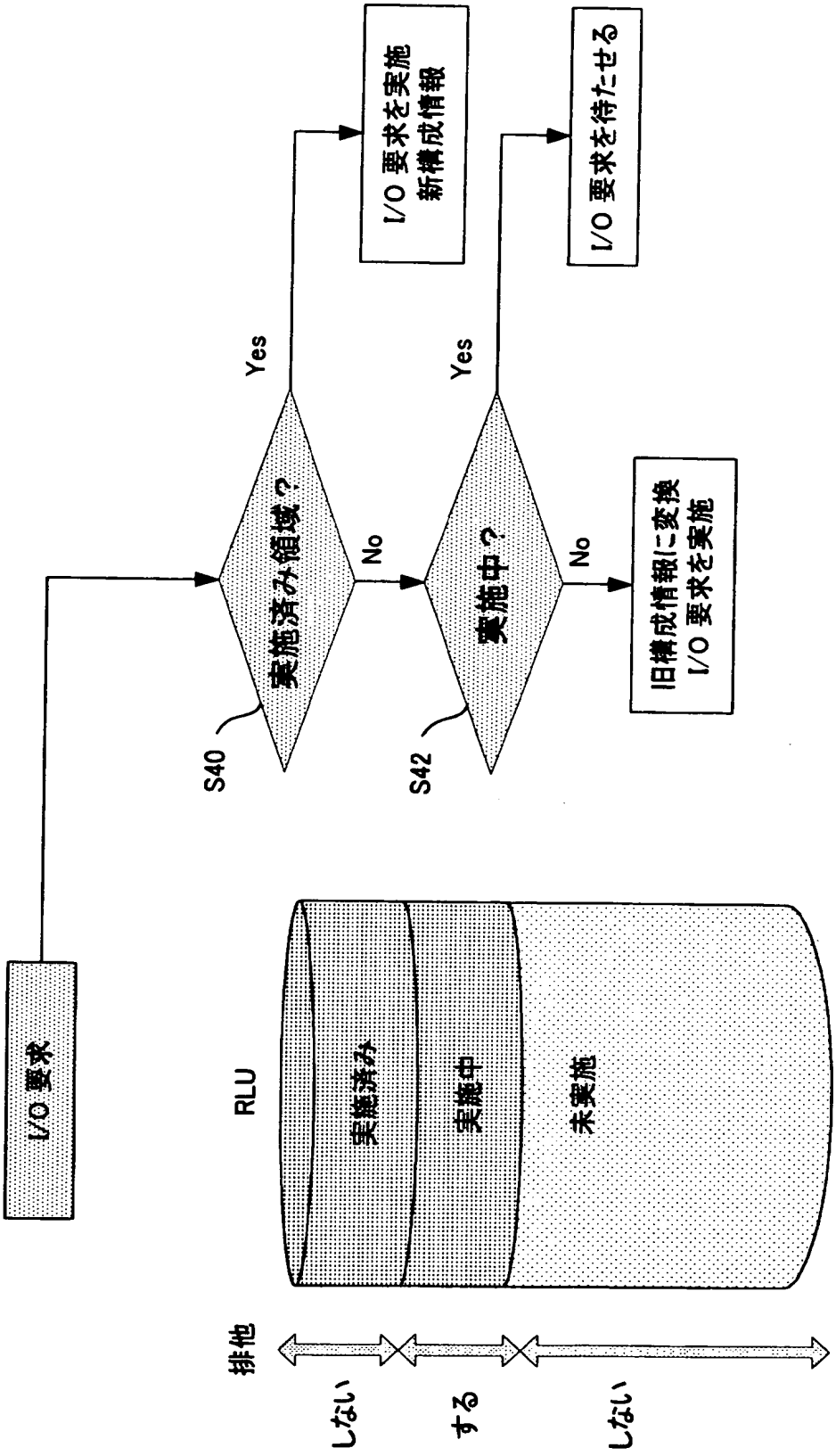
【図１８】


一回の処理サイズ
（仕掛け中大きさ）
獲得可能なサイズ？
Yes
No
一回だけ要求
複数回獲得？
Yes
No
必要回数を要求
一回だけ要求し、使いまわすようにする

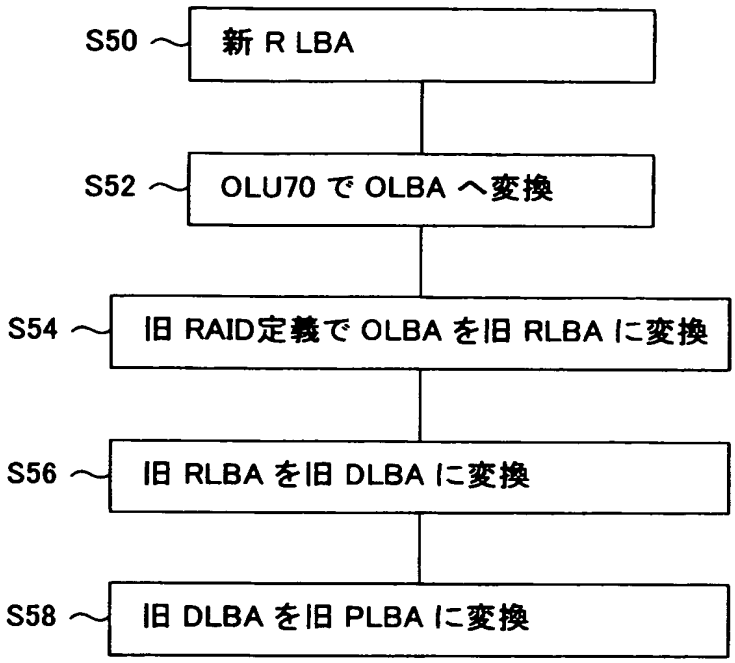
【図１９】


Disk#
00
01
02
03
09
Stripe
00
01
02
03
04
新構成 (Expansion 後)
1/2回目
2/2回目
旧構成 (Expansion 前)

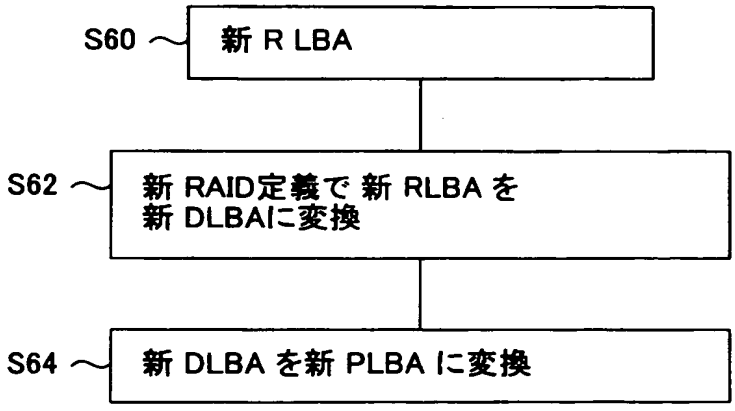
【図２０】


I/O 要求
RLU
実施済み
実施中
未実施
排他
しない
する
しない
S40
実施済み領域？
Yes
No
I/O 要求を実施
新構成情報
S42
実施中？
Yes
No
I/O 要求を待たせる
旧構成情報に変換
I/O 要求を実施

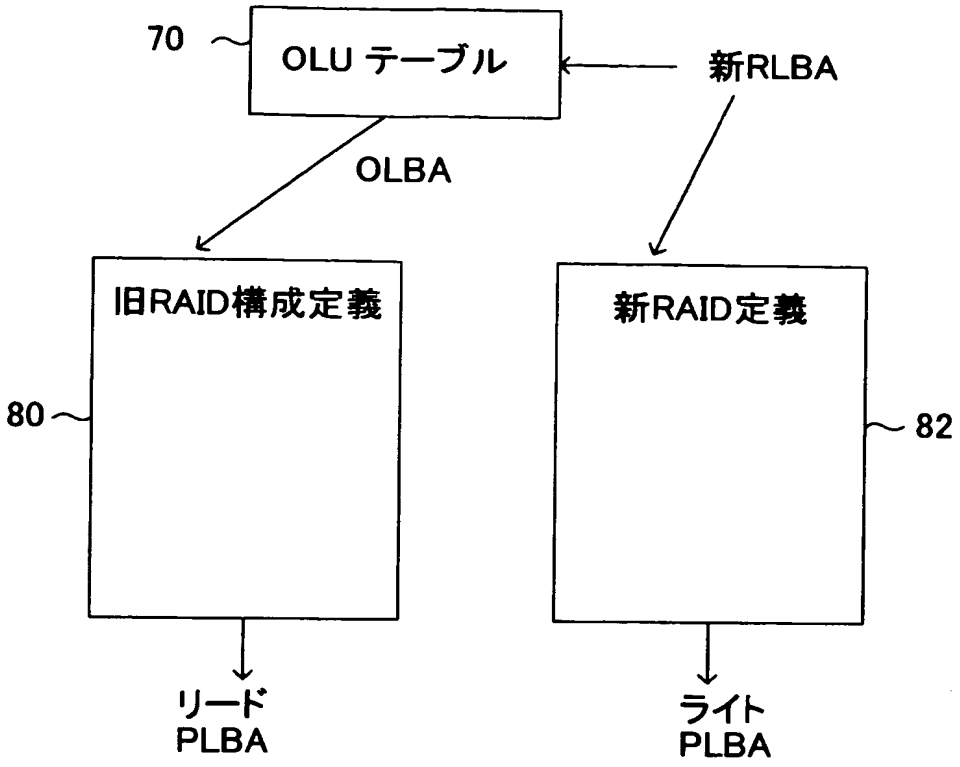
【図２１】


S50
新 R LBA
S52
OLU70 で OLBA へ変換
S54
旧 RAID定義で OLBA を旧 RLBA に変換
S56
旧 RLBA を旧 DLBA に変換
S58
旧 DLBA を旧 PLBA に変換

【図２２】


S60
新 R LBA
S62
新 RAID定義で 新 RLBA を
新 DLBAに変換
S64
新 DLBA を新 PLBA に変換

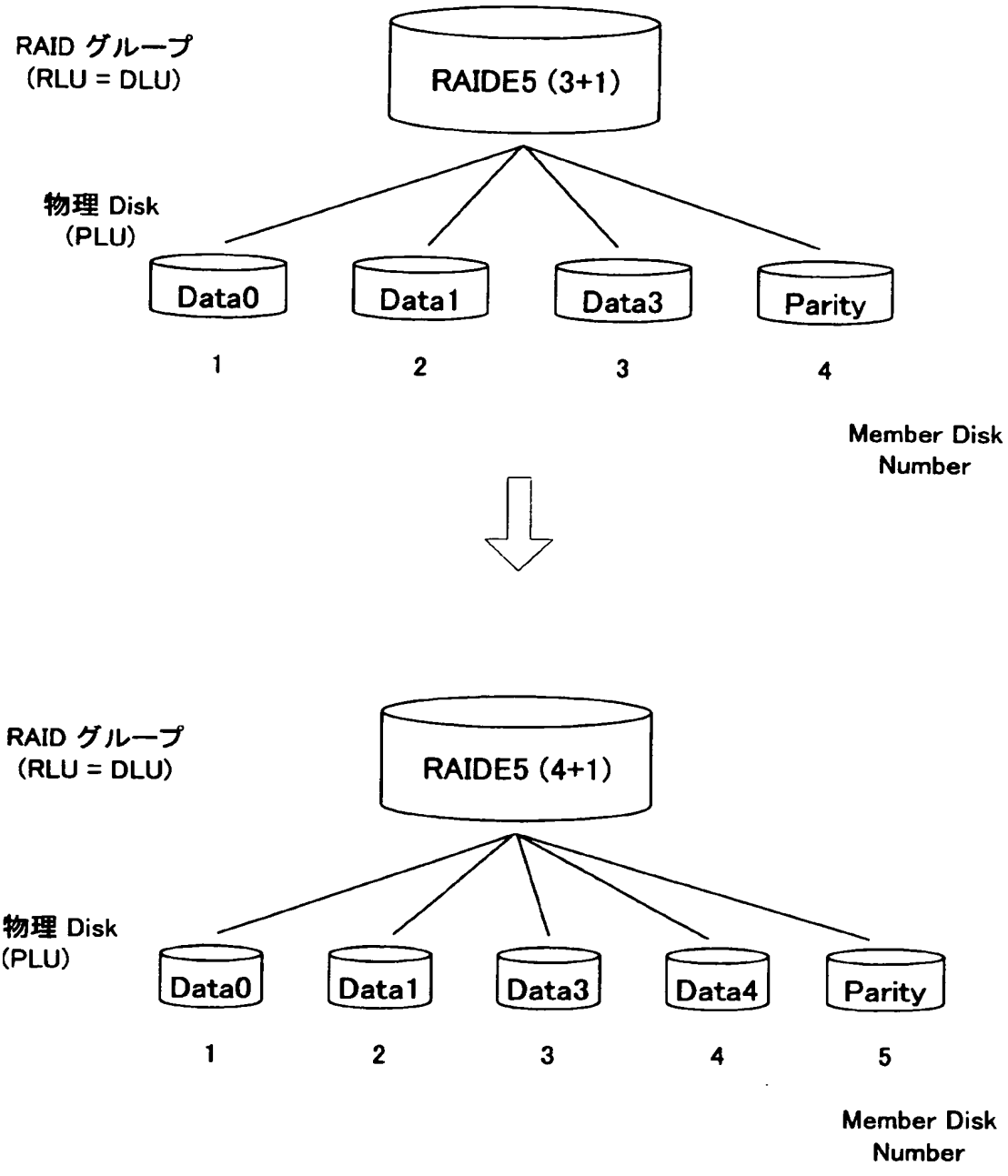
【図２３】


70
OLU テーブル
新RLBA
OLBA
旧RAID構成定義
80
新RAID定義
82
リード
PLBA
ライト
PLBA

【図２４】

	構成定義情報		LDE状況			
	LUN	項目	実施前	処理中	終了	
RAID定義	RLU	RLUN	0	←	←	
		Level	RAID5	←	←	
		MemberDisks	4	5	←	
		PLUNs	10-13	10-14	←	
		StartLBA	0	←	←	
		BlockCount	400	500	←	~82
		CheckCode	id#0	←	←	
		Status (VM mode)	0	LDE Flag	0	
	DLU	DLUN	0	←	←	
	PLU					
テンポラリRAID定義	RLU	RLUN	-	255	-	
		Level	-	RAID5	-	
		MemberDisks	-	4	-	
		PLUNs	-	10-13	-	~80
		StartLBA	-	0	-	
		BlockCount	-	400	-	
		CheckCode	-	id#0	-	
		Status (VM mode)	-	T Flag	-	
	DLU	DLUN	-	255	-	

【図２５】


RAID グループ
(RLU = DLU)
RAIDE5 (3+1)
物理 Disk
(PLU)
Data0
Data1
Data3
Parity
1
2
3
4
Member Disk
Number
RAID グループ
(RLU = DLU)
RAIDE5 (4+1)
物理 Disk
(PLU)
Data0
Data1
Data3
Data4
Parity
1
2
3
4
5
Member Disk
Number

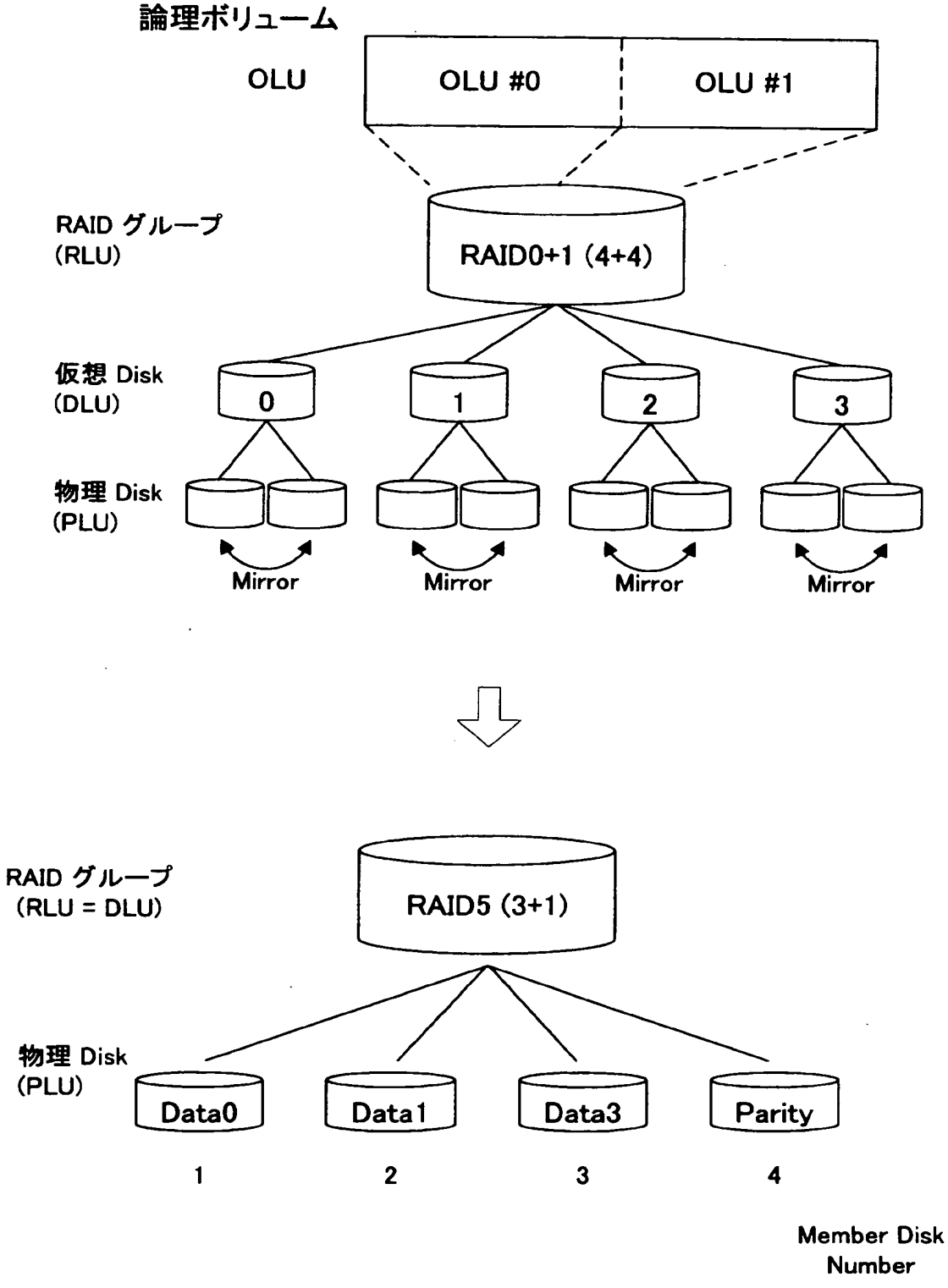
【図２６】

	構成定義情報		LDE状況		
	LUN	項目	実施前	処理中	終了
RAID定義	RLU	RLUN	0	←	←
		Level	RAID0+1	RAID5	←
		MemberDisks		4	←
		PLUNs		0～3	←
		StartLBA	0	←	←
		BlockCount	400		←
		CheckCode	id#0	←	←
		CVM mode	0	LDE Flag	0
	DLU	DLUN	0	←	←
	PLU				
テンポラリRAID定義	RLU	RLUN	-	255	-
		Level	-	RAID0+1	-
		MemberDisks	-	4	-
		PLUNs	-	0～3	-
		StartLBA	-	0	-
		BlockCount	-	400	-
		CheckCode	-	id#0	-
		CVM mode	-	T Flag	-
	DLU	DLUN	-	255	-
	PLU	-	-	-	-

82

80

【図２７】

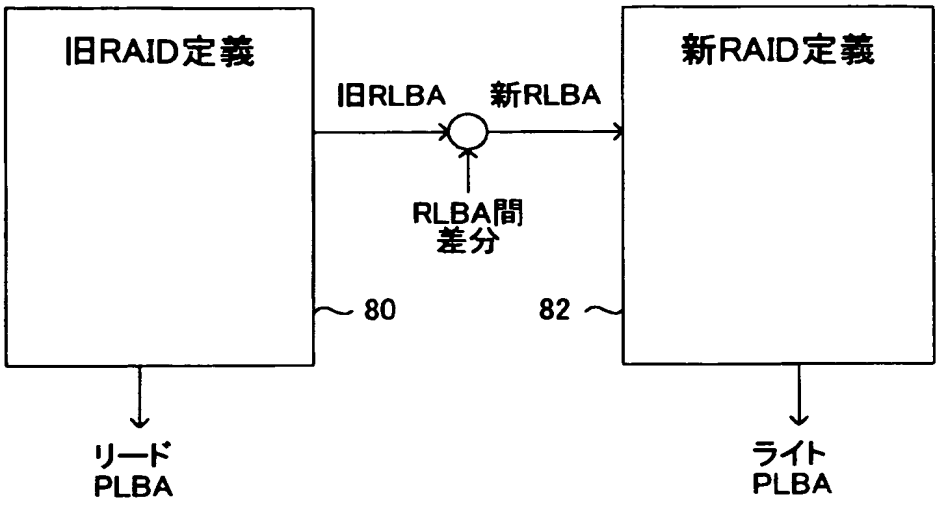

論理ボリューム
OLU
OLU #0
OLU #1
RAID グループ
(RLU)
RAID0+1 (4+4)
仮想 Disk
(DLU)
0
1
2
3
物理 Disk
(PLU)
Mirror
Mirror
Mirror
Mirror
RAID グループ
(RLU = DLU)
RAID5 (3+1)
物理 Disk
(PLU)
Data0
Data1
Data3
Parity
1
2
3
4
Member Disk
Number

【図２８】

RLU		DLU 1	
RAID Level	0+1	Stripe Depth	128
Stripe Depth	128	Stripe Size	128
Stripe Size	128x4	Member Disk Count	2
Member Disk Count	4	DLU 2	
		DLU 3	
		DLU 4	

RLU		DLU	
RAID Level	5	Stripe Depth	128
Stripe Depth	128x3	Stripe Size	128x3
Stripe Size	128x3	Member Disk Count	4
Member Disk Count	1		

【図２９】



【書類名】　要約書

【要約】

【課題】ＲＡＩＤ構成の冗長度を変更するＲＡＩＤ装置において、多様なＲＡＩＤレベル変換、容量増加を可能とする。

【解決手段】少なくとも、ＲＡＩＤレベルと論理デバイス数を定義した新旧のＲＡＩＤ構成定義情報（８０、８１）を使用し、制御部（１０）が、それぞれによりＲＬＵマッピングして、ステージング、ライトバックして、ＲＡＩＤ構成を変更する。このため、多様なＲＡＩＤレベルの変換、容量増加を実現できる。

【選択図】図１５

特願２００２－３７８２８４

出　願　人　履　歴　情　報

識別番号　　［０００００５２２３］

１．変更年月日　　１９９６年　３月２６日
［変更理由］　　住所変更
住　所　　神奈川県川崎市中原区上小田中４丁目１番１号
氏　名　　富士通株式会社

特願２００２－３７８２８４

出　願　人　履　歴　情　報

識別番号　　［０００１３６１３６］

１．変更年月日　　１９９０年　８月３１日
［変更理由］　新規登録
住　所　石川県河北郡宇ノ気町字宇野気ヌ９８番地の２
氏　名　株式会社ピーエフユー

２．変更年月日　　２００３年　４月　７日
［変更理由］　名称変更
住　所　石川県河北郡宇ノ気町字宇野気ヌ９８番地の２
氏　名　株式会社ＰＦＵ